Approximate Derivative Pricing for Large Classes of Homogeneous Assets with Systematic Risk*

Patrick Gagliardini†

University of Lugano and Swiss Finance Institute

Christian Gouriéroux

CREST (Paris), CEPREMAP and University of Toronto

First Version: March 2008

This Version: June 2010

*Acknowledgements: We thank D. Bates, F. Galli, M. Gordy, A. Monfort, the co-editor E. Renault, two anonymous referees and the participants at the International Financial Risk Forum 2008 in Paris, the FINRISK Research Day 2008 in Gerzensee, the Credit Conference 2008 in Venice, ESEM 2009 in Milan and the SoFiE 2009 Conference in Chicago for very useful comments. The first author gratefully acknowledges financial support of the Swiss National Science Foundation through the NCCR FINRISK network. The second author gratefully acknowledges financial support of NSERC Canada and the chair AXA/Risk Foundation: "Large Risks in Insurance".

†Corresponding author: Patrick Gagliardini, Faculty of Economics, University of Lugano, Via Buffi 13, CH-6900 Lugano, Switzerland. Phone: +41 58 666 4660. Email: patrick.gagliardini@usi.ch.

Approximate Derivative Pricing for Large Classes of Homogeneous Assets with Systematic Risk

Abstract

We consider a homogeneous class of assets, whose returns are driven by an unobservable factor representing systematic risk. We derive approximated pricing formulas for the future factor values and their proxies, when the size n of the class is large. Up to order $1/n$, these closed form approximations involve well-chosen summary statistics of the basic asset returns, but not the current and lagged factor values. The potential of the closed form approximation formulas seems quite large, especially for credit risk analysis, which considers large portfolios of individual loans or corporate bonds, and for longevity risk analysis, which involves large portfolios of life insurance contracts.

Keywords: Derivative Pricing, Large Portfolio, Systematic Risk, Credit Risk, Basket Derivatives, Credit Default Swap, Default Correlation, Granularity Adjustment.

JEL classification: G12, C23.

The interest in homogeneous classes of assets is growing rapidly with the Basel 2 [BCBS (2001)] and Solvency 2 [CEA (2007)] regulations and the introduction of basket, or factor derivatives. The portfolios involved in the balance sheet of a bank or an insurance company can involve several millions of individual assets, and it is suggested to cluster them into homogeneous classes to facilitate the computation of risk measures. Examples of large homogeneous classes of assets are:

(i) a set of corporate bonds for firms with given industrial sector, rating and country;

(ii) a cohort of mortgages with identical contractual interest rate, maturity, pattern of monthly payment and identical borrower's rating;

(iii) a set of life insurance contracts with similar design and contractors with identical age.

The risk analysis in such large homogeneous classes of assets relies on the distinction between systematic risks, which arise from assets' exposure to common factors, and unsystematic risks, which are specific to the individual assets. For instance, a stationary common factor is usually introduced in examples (i) and (ii) above to capture the default correlation. Among others, Vasicek (1987), Gupton et al. (1997), Schoenbucher (2001), Hull, White (2004), Laurent, Gregory (2005) consider a static framework, while Duffie et al (2009) consider a dynamic framework with observable individual specific variables as well as an unobservable common factor, called frailty. A non-stationary common factor is usually introduced in example (iii) to capture the longevity risk associated with the increase in life expectancy [see e.g. Lee, Carter (1992), Dahl, Moller (2004), Cairns, Blake, Dowd (2006), Schrager (2006), Gouriéroux, Monfort (2008)]. Basket derivatives, such as basket default swaps (BDS), tranches of Collateralized Debt Obligations (CDO), stripped components of Mortgage Backed Securities (MBS) and mortality linked securities (MLS) have been introduced to hedge the systematic risk factor [1]. For this purpose, these derivatives are written on proxies of the unobservable factor.

When pricing basket derivatives, two important issues have to be taken into account:

(i) The markets are incomplete due to the discrete time framework, the possibility of default and the effect of (unobservable) dynamic factors. Therefore, there is a multiplicity of risk-neutral distributions.

(ii) Even if some specific form of the risk-neutral distribution is selected, the derivative prices depend on the current (and lagged) factor values. In the literature, it is usually assumed that these unobservable factor values can only be recovered by means of observed prices of derivatives traded on the market.

However, if such derivatives are not actively traded, the factor values cannot be deduced and the pricing formula cannot be applied. This drawback arises typically at the emergence of new derivative markets, when a coherent quotation scheme for derivatives not yet highly traded has to be proposed. This paper explains how this drawback for the first quotation of derivatives can be circumvented when we observe the prices (or returns) of a large homogeneous class of assets driven by a same dynamic factor. We show that the unobserved factor value and the derivative prices can be approximated by means of the observed asset prices (returns) at order $1/n$, where n is the size of the class. Thus, it is not necessary to observe at least one additional derivative price to be able to price coherently the other ones. By exploiting the large class size, our approximate derivative pricing formulas are based on a closed form Gaussian approximation of the predictive distribution of the factor value. This sharply differs from other filtering approaches proposed for dynamic latent factor models in the default risk literature [see e.g. Duffie et al (2009)], where the predictive distribution of the factor is computed with simulation based methods.

The notion of a homogeneous class of risks is defined in Section 1 by means of the conditional historical and risk-neutral distributions of risk variables $y_{1,t}, ..., y_{n,t}$, say, given the underlying path of the factor f_t, and the historical and risk-neutral factor dynamics. We particularize this notion to different types of risk to account for quantitative as well as qualitative risk variables. Section 2 provides a convenient closed form approximation of the historical and risk-neutral distributions of a future factor value $f_{t+\tau}$ at time-to-maturity $\tau \geq 1$ given the information available at time t, which includes the current and past values of $y_{1,t}, ..., y_{n,t}$, for large n. We discuss the consequences of this approximation formula in terms of first quotation of derivatives. In Section 3, the methodology is illustrated by focussing on the approximate pricing of derivatives written on a common factor proxy. A numerical illustration to the pricing of basket default swaps is given in Section 4. Section 5 concludes. The proofs are gathered in the Appendices.

1 Homogeneous classes of assets

Let us consider n assets with risk characteristics observed at T different dates. The risk characteristics are denoted y_{it}, for $i = 1, ..., n$, and $t = 1, ..., T$, and have different interpretations and ranges according to the assets. For a class of stocks, y_{it} can be the return between $t-1$ and t, and is real valued. For a class of corporate bonds, $y_{i,t}$ can be the issuer default indicator, that is equal to 1 if the issuer of bond i is in default at date t, and 0, otherwise. Alternatively, y_{it} can be the spread, that is the difference between the corporate and risk-free interest rates, and is non-negative. Finally, for a class of digital Credit Default Swaps (CDS) with given time-to-maturity, y_{it} can be the ratio of the CDS price to the price of the risk-free zero-coupon bond with the same maturity. Then, y_{it} takes value between 0 and 1 and corresponds to a risk-neutral (implied) default probability.

1.1 Definition and assumptions

The notion of homogeneous class is defined for instance in Gouriéroux, Tiomo (2007), Chapter 7.

Definition 1: *A class of assets is homogeneous under the historical probability (resp. the risk-neutral probability), if and only if the joint historical (resp. risk-neutral) distribution of processes* $(y_{1,t}), ..., (y_{n,t})$ *is invariant by permutation* [2].

As remarked in the Introduction, the homogeneity assumption underlies the analysis by segment recommended in the current regulation. This explains the focus of the paper on homogenous classes of assets. However, the results can be easily extended to observed heterogeneity [see Gagliardini, Gouriéroux, Monfort (2010)]. Moreover, the exchangeability property of the assets in Definition 1 can be equivalently written in terms of underlying factors by applying de Finetti's Theorem [de Finetti (1931)] and its generalization by Hewitt and Savage (1955). For expository purpose, and since basket derivatives are typically introduced to hedge a given single systematic risk, we focus on the single-factor case. However, the results can be extended to the multifactor framework.

Assumption A.1: *Under the historical probability, the processes* $(y_{1,t}), ..., (y_{n,t})$ *are independent given the factor path* (f_t).

Assumption A.2: *The conditional historical density of* $y_{1,t},...,y_{n,t}$ *given the past values of the* $y_{i,t}$*'s, and the current and past values of the factor, is* $\prod_{i=1}^{n} h\left(y_{it}|f_t\right)$*, say, w.r.t. some dominating measure* $\prod_{i=1}^{n} dy_{i,t}$. [3]

The conditional historical density is driven by the current factor value only, is independent of the date, and the variables $y_{1,t},...,y_{n,t}$ are i.i.d. conditional on the current factor value. This explains the terminology "conditionally independent risk model" introduced in the credit risk literature [see e.g. Schoenbucher (2001)]. The conditional density $\prod_{i=1}^{n} h\left(y_{it}|f_t\right)$ characterizes the contemporaneous effect of factor f_t on the individual risks and is called micro-density. Factor f_t represents systematic risk [4]. Further, the dynamics of variables $y_{i,t}$, $i = 1,...,n$, is through the dynamics of factor f_t only, which is specified next.

Assumption A.3: *Under the historical probability, the factor process is Markovian with transition density* $g\left(f_t|f_{t-1}\right)$*, say, w.r.t. a dominating measure* df_t.

Under Assumptions A.1, A.2 and A.3, the joint conditional distribution of $y_{1,t+1},...,y_{n,t+1},f_{t+1}$ given the current and past values of the $y_{i,t}$'s and f_t is:

$$l(y_{1,t+1},...,y_{n,t+1},f_{t+1}|\underline{y_{1,t}},...,\underline{y_{n,t}},\underline{f_t}) = g\left(f_{t+1}|f_t\right)\prod_{i=1}^{n} h\left(y_{i,t+1}|f_{t+1}\right), \tag{1.1}$$

where $\underline{y_{i,t}}$ denotes the individual history $y_{i,t},y_{i,t-1},...$, and similarly for $\underline{f_t}$. It is easily checked from equation (1.1) that the assets are exchangeable under the historical probability. Assumptions A.1-A.3 correspond to a nonlinear state space model, where the measurement equations are characterized by the conditional density $\prod_{i=1}^{n} h\left(y_{it}|f_t\right)$ and the state equation by the transition density $g(f_t|f_{t-1})$.

The above set of assumptions is completed by assumptions on the investors' information set and on the stochastic discount factor (sdf) $m_{t,t+1}$, which characterizes the change of measure to pass from the historical probability to the pricing operator [see e.g. Harrison, Kreps (1979), Hansen, Richard (1987)].

Assumption A.4: *The investors' information set at date t is* $\Omega_t = \left(\underline{y_{1,t}},...,\underline{y_{n,t}},\underline{f_{t-1}}\right)$.

Assumption A.5: *The sdf between date* t *and date* $t+1$ *is* $m_{t,t+1} = m(f_t)$.

The information set of the investors includes the current and past values of the $y_{i,t}$'s, and the past values of the factor. Hence, the variables $y_{i,t}$, $i = 1, ..., n$, are observed at the beginning of period $(t, t+1)$, whereas f_t is observed by the investor at the end of that period. Assumption A.4 reflects the incompleteness of the investors' information, since the current factor value f_t is unobservable for the investor at date t. Thus, the investor faces a filtering problem when pricing factor derivatives. In continuous time, it would not be possible to distinguish between the current factor value f_t and the most recent lagged value f_{t^-}, say, for a factor with continuous path. In discrete time, the most recent factor value f_{t-1} differs from f_t, and Assumption A.4 on information becomes relevant.

The current factor value unobservable at time t is included in the sdf $m_{t,t+1}$, which is really stochastic from the investor's view point at date t and is used for risk correction and time discount between t and $t+1$. The existence of a sdf is a consequence of the no arbitrage assumption. In general, the sdf between t and $t+1$ depends on the investors information at $t+1$, that is $\underline{y_{1,t+1}}, ..., \underline{y_{n,t+1}}, \underline{f_t}$, and on the market, in particular on the number n of assets. Under Assumption A.5, the sdf is supposed to depend only on the current value of the common risk factor, and neither on the idiosyncratic risks specific of the individual assets, nor on the size n of the homogeneous class. The assumption that the sdf does not depend on the idiosyncratic risks is standard for pricing credit securities [see e.g. Gouriéroux, Monfort, Polimenis (2006)] and mortality linked securities (MLS) [see e.g. Schrager (2006)]. This assumption is justified by the underlying point of view that idiosyncratic risks can be diversified and as such their prices are not adjusted for risk, similarly as in the standard arbitrage pricing theory [Ross (1976)]. The assumption of sdf independence w.r.t. size n merits also to be further discussed. Let us assume an infinite population, $n = \infty$. Then, f_t becomes known (by using a cross-sectional maximum likelihood estimator, see Section 2) and the short-term interest rate $r_t = -\log E\left[m_{t,t+1}|\Omega_t\right] = -\log m(f_t)$ is known too. Thus, either both function m and the interest rate are constant, or factor f_t can be identified with the interest rate, up to a given transformation. If we want to get a pricing model with constant interest rate r and non constant factor in the limit $n = \infty$, it is necessary to suppose $m_{t,t+1} = m(n, f_t)$, with $m(\infty, f_t) = \exp(-r)$, and if we want to get a stochastic interest rate r_t and another stochastic factor f_t to suppose $m_{t,t+1} = m(r_t, n, f_t)$, with $m(r_t, \infty, f_t) = \exp(-r_t)$. A theoretical justification of the sdf specification would require a structural equilibrium model, which is beyond the scope of this paper.

Under Assumptions A.1-A.5, the short-term pricing kernel w.r.t. the dominating measure $df_t \prod_{i=1}^{n} dy_{i,t+1}$ is:

$$p(y_{1,t+1}, ..., y_{n,t+1}, f_t|\Omega_t) = m(f_t) g(f_t|\Omega_t) \int \prod_{i=1}^{n} h(y_{i,t+1}|f_{t+1})\, g(f_{t+1}|f_t)\, df_{t+1}, \tag{1.2}$$

where:

$$g(f_t|\Omega_t) = \frac{g(f_t|f_{t-1}) \prod_{i=1}^{n} h(y_{i,t}|f_t)}{\int g(f_t|f_{t-1}) \prod_{i=1}^{n} h(y_{i,t}|f_t)\, df_t}, \tag{1.3}$$

denotes the density of f_t given the investor information Ω_t at t. This pricing kernel is used to compute the price at date t of any short-term derivative written on $y_{1,t+1}, ..., y_{n,t+1}, f_t$ with payoff $a(y_{1,t+1}, ..., y_{n,t+1}, f_t)$ as:

$$\pi_t(a, 1) = \int a(y_{1,t+1}, ..., y_{n,t+1}, f_t) p(y_{1,t+1}, ..., y_{n,t+1}, f_t|\Omega_t)\, dy_{1,t+1}...dy_{n,t+1} df_t.$$

The pricing kernel depends on information Ω_t through the current observations $y_{1,t}$, ..., $y_{n,t}$ and the past factor value f_{t-1} only.

Proposition 1. *The historical and risk-neutral conditional distributions of $y_{1,t+1}, ..., y_{n,t+1}$ given f_t are the same. The change of measure between the historical and risk-neutral distributions of f_t conditional on Ω_t is given by the sdf $m(f_t)$.*

Proof. From the short-term pricing kernel (1.2), we deduce the short-term risk-neutral distribution by standardizing with the short-term zero-coupon bond price $B(t, t+1) = \int m(f_t) g(f_t|\Omega_t)\, df_t$. Thus, the risk-neutral conditional distribution is:

$$\begin{aligned} l^*(y_{1,t+1}, ..., y_{n,t+1}, f_t|\Omega_t) &= p(y_{1,t+1}, ..., y_{n,t+1}, f_t|\Omega_t) / B(t, t+1) \\ &= g^*(f_t|\Omega_t) \int \prod_{i=1}^{n} h(y_{i,t+1}|f_{t+1})\, g(f_{t+1}|f_t)\, df_{t+1}, \end{aligned}$$

where:

$$g^*\left(f_t|\Omega_t\right) = \frac{m(f_t)g\left(f_t|\Omega_t\right)}{\displaystyle\int m(f_t)g\left(f_t|\Omega_t\right)df_t},$$

is the risk-neutral conditional density of f_t given Ω_t. The result follows. □

1.2 Examples

Let us now derive the expression of the micro-density $\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right)$ for standard models encountered in the literature for homogeneous class of stocks, corporate bonds, or digital CDS. In general, the factor has a nonlinear effect and may admit different interpretations. For instance, it can be a stochastic mean, a stochastic variance, a stochastic default probability, or a stochastic concentration parameter.

i) Linear factor model

Let us consider a linear factor model:

$$y_{it} = a + bF_t + \sigma u_{it}, \quad i = 1, ..., n,$$

where the errors u_{it} are independent standard Gaussian variables. The common factor F_t impacts the conditional mean of the variables y_{it}. When the variables y_{it}, $i = 1, ..., n$, are stock returns, and factor F_t is the market portfolio return, we get the standard market model written for a homogeneous class of stocks, since the alphas, betas and idiosyncratic volatilities are stock independent. The conditional correlation between any two assets is zero, but the unconditional correlation is non-zero when the unobservable factor F_t is integrated out. By introducing the transformed factor $f_t = a + bF_t$, this model satisfies Assumptions A.1 and A.2 with micro-density:

$$\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) = \frac{1}{\left(2\pi\sigma^2\right)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_{it}-f_t\right)^2\right\}.$$

ii) Single-factor stochastic volatility model

Let us consider the model:

$$y_{it} = \mu + f_t^{1/2} u_{it}, \quad i = 1, ..., n,$$

where factor (f_t) is a positive Markov process, and the errors u_{it} are independent standard Gaussian variables. Factor (f_t) introduces a dependence between realized volatilities computed on the individual assets. The joint stochastic volatility-covolatility matrix of the n assets is $\Sigma_t = f_t Id_n$, where Id_n denotes the identity matrix of order n. The associated micro-density is:

$$\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{n}{2}\log f_t - \frac{1}{2f_t}\sum_{i=1}^{n}\left(y_{it}-\mu\right)^2\right\}.$$

iii) Single Risk Factor (SRF) model for default correlation

The Single Risk Factor (SRF) model is specified as [see Vasicek (1987) and the Credit Metrics framework in Gupton, Finger, Bhatia (1997)]:

$$\log\left(A_{it}/L_{it}\right) = \mu + \sigma\sqrt{\rho}F_t + \sigma\sqrt{1-\rho}u_{it}, \tag{1.4}$$

where the error terms $u_{i,t}$ are independent standard Gaussian variables, A_{it} and L_{it} denote the asset value and liability of firm i at date t, respectively, and ρ is a parameter [5] in $(0, 1)$. When the variance of the common factor F_t is normalized to 1, the unconditional variance of the log asset-to-liability ratio of any firm is σ^2, and the unconditional correlation between the log asset-to-liability ratios of any two firms is ρ. This structural model is used to characterize the joint distribution of default occurrence: $y_{it} = 1$, if $A_{it} < L_{it}$, and $= 0$, otherwise. Conditional on factor value F_t, the dichotomous variables $y_{i,t}$, $i = 1, ..., n$, are i.i.d. with Bernoulli distribution $\mathcal{B}\left(1, f_t\right)$, where the canonical factor is defined by $f_t = \Phi\left(-\frac{\mu+\sigma\sqrt{\rho}F_t}{\sigma\sqrt{1-\rho}}\right)$ and Φ denotes the cumulative distribution function of a standard Gaussian variable. This transformed factor f_t is the conditional default probability at time t. The joint conditional distribution of default indicators is:

$$\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) = \left(f_t\right)^{n\bar{y}_{nt}}\left(1-f_t\right)^{n\left(1-\bar{y}_{nt}\right)}, \tag{1.5}$$

where $\bar{y}_{nt} = \frac{1}{n}\sum_{i=1}^{n} y_{i,t}$ is the proportion of obligors that defaulted in period t.

iv) Single-factor corporate spread model

Corporate spreads are positive and their dynamics is generally specified to get an affine term structure [see e.g. Duffie, Filipovic, Schachermayer (2003) and Duffie, Singleton (2003) in continuous time, Darolles, Gouriéroux, Jasiak (2006) and Gouriéroux (2008) in discrete time]. For instance, let us assume that the historical conditional distribution of y_{it} given f_t is a gamma distribution $\gamma\left(f_t, \lambda\right)$ with stochastic degree of freedom f_t and scale parameter λ. The conditional density is:

$$h\left(y_{i,t}|f_t\right) = \frac{1}{\Gamma(f_t)} \exp\left(-\lambda y_{i,t}\right) y_{i,t}^{f_t-1} \lambda^{f_t} \; \mathbb{1}_{y_{i,t}>0},$$

and the micro-density becomes:

$$\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) = \frac{1}{\Gamma(f_t)^n} \exp\left(-\lambda \sum_{i=1}^{n} y_{i,t}\right) \left(\prod_{i=1}^{n} y_{i,t}\right)^{f_t-1} \lambda^{n f_t} \; \mathbb{1}_{\min_i y_{i,t}>0},$$

where Γ denotes the gamma function. The associated conditional Laplace transform is given by:

$$E\left[\exp\left(-u y_{i,t}\right)|f_t\right] = \exp\left[f_t \log(1 + u/\lambda)\right].$$

This Laplace transform is an exponential affine function of the factor, which simplifies the derivation of nonlinear predictions if process (f_t) is affine [6].

v) Homogeneous class of CDS

In this example, the variable y_{it} is the value of a digital CDS [7] divided by the associated zero-coupon bond price with identical time-to-maturity. Such variable takes continuous values between 0 and 1, and its distribution can be chosen in the class of beta distributions:

$$h\left(y_{it}|f_t\right) = \frac{\Gamma(f_t)}{\Gamma(\delta f_t)\Gamma\left[(1-\delta) f_t\right]} y_{it}^{\delta f_t - 1} \left(1 - y_{it}\right)^{(1-\delta) f_t - 1} \; \mathbb{1}_{0<y_{it}<1}, \tag{1.6}$$

where δ is a scalar parameter in $(0,1)$, and f_t is a positive factor. The conditional mean $E\left(y_{it}|f_t\right) = \delta$ is constant. Moreover, the conditional variance of a variable on $[0,1]$ is upper bounded:

$$V\left(y_{it}|f_t\right) \leq E\left(y_{it}|f_t\right)\left[1 - E\left(y_{it}|f_t\right)\right] = \delta(1-\delta),$$

and the upper bound is reached when the total mass is on the two-points set $\{0,1\}$. It is easily checked from (1.6) that:

$$f_t + 1 = \frac{\delta(1-\delta)}{V\left(y_{it}|f_t\right)}.$$

Thus, factor f_t measures the concentration of the distribution, taking into account the existence of the upper bound. We get a model with constant conditional mean and stochastic concentration parameter. The micro-density is:

$$\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) = \left(\frac{\Gamma(f_t)}{\Gamma(\delta f_t)\Gamma\left[(1-\delta)\, f_t\right]}\right)^{n} \left(\prod_{i=1}^{n} y_{i,t}\right)^{\delta f_t - 1} \left(\prod_{i=1}^{n} (1-y_{i,t})\right)^{(1-\delta) f_t - 1} \prod_{i=1}^{n} \mathbb{1}_{0<y_{it}<1}.$$

By interchanging the roles of δ and f_t in density (1.6), we get an alternative single-factor model with stochastic conditional mean f_t and constant concentration parameter δ.

2 Large portfolio approximation

We first derive the approximation theorems of the predictive factor distribution valid for large n. Then, we explain how these theorems are used for pricing purposes by applying them to the risk-neutral distribution.

2.1 Approximation theorems

Let us consider a homogeneous class satisfying Assumptions A.1-A.3. The predictive distribution of factor f_t is characterized by its Laplace transform, which gives the conditional expectation of any exponential transformation of the factor given the investors' information.

Assumption A.6: *The conditional Laplace transform of f_t given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}$, that is, the function:*

$$\mathcal{L}_{n,t}(u) = E\left[\exp\left(u f_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right], \quad u \in \mathbb{R},$$

is well-defined in a neighbourhood of $u = 0$.

The approximation theorem is derived along the lines of the Laplace method [see e.g. Jensen

(1995)]. Let us define the cross-sectional maximum likelihood approximation of the factor:

$$\hat{f}_{nt} = \arg\max_{f_t} \sum_{i=1}^{n} \log h\left(y_{it}|f_t\right), \tag{2.1}$$

and introduce the next Assumption A.7.

Assumption A.7: *The function* $f \to E[\log h(y_{i,t}|f)|f_t]$ *is uniquely maximized at* $f = f_t$*, for any* t*,* P*-a.s. Moreover,* $I_t := E\left[-\frac{\partial^2 \log h(y_{i,t}|f_t)}{\partial f_t^2}|f_t\right] > 0$*, for any* t*,* P*-a.s.*

Assumption A.7 corresponds to the standard global and local identification conditions in maximum likelihood estimation, when the unobservable factor value at date t is treated as an unknown parameter for the cross-section at date t. Under Assumption A.7, the cross-sectional log-likelihood function $\sum_{i=1}^{n} \log h\left(y_{it}|f_t\right)$ admits an unique global maximum w.r.t. f_t, that is $\hat{f}_{n,t}$, on any compact interval, with probability approaching 1 as the cross-sectional sample size n goes to infinity.

Proposition 2. *Under Assumption A.1-A.7, the conditional Laplace transform of* f_t *given* $\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}$ *is such that:*

$$\begin{aligned}
\mathcal{L}_{n,t}(u) &= \exp\left\{u\left(\hat{f}_{nt} + \frac{1}{n} I_{nt}^{-1} \frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|\hat{f}_{n,t-1}\right) + \frac{1}{2}\frac{1}{n} I_{nt}^{-2} K_{nt}\right) + \frac{1}{2}\frac{1}{n} I_{nt}^{-1} u^2 + o(1/n)\right\} \\
&= E\left[\exp\left(u f_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}\right] + o(1/n),
\end{aligned}$$

where:

$$I_{nt} := -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \log h}{\partial f_t^2}\left(y_{it}|\hat{f}_{nt}\right) \quad \textit{and} \quad K_{nt} := \frac{1}{n}\sum_{i=1}^{n} \frac{\partial^3 \log h}{\partial f_t^3}\left(y_{it}|\hat{f}_{nt}\right).$$

Proof. See Appendix 1. □

Up to order $1/n$, the dependence of $\mathcal{L}_{n,t}(u)$ on information is captured by means of the four summary statistics $\hat{f}_{nt}$, $\hat{f}_{n,t-1}$, I_{nt} and K_{nt}, which depend only on current and past values $y_{i,t}$. **Thus, the lagged factor values are non-informative for large class size (at order** $1/n$**).** The statistic $\hat{f}_{nt}$ is the cross-sectional Maximum Likelihood (ML) estimator of the current factor value f_t. Statistic I_{nt} is the estimated Fisher information corresponding to this estimation problem. Finally, K_{nt} is an additional statistic involved in asymptotic bias correction. This set of statistics is independent of the factor dynamics, that is, of function g. By Proposition 2, the logarithm of the approximated Laplace transform is quadratic in argument u; thus, the conditional distribution of f_t given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}$ is approximately normal at order $1/n$:

$$N\left(\hat{f}_{nt} + \frac{1}{n}\left[I_{nt}^{-1}\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|\hat{f}_{n,t-1}\right) + \frac{1}{2}I_{nt}^{-2}K_{nt}\right], \frac{1}{n}I_{nt}^{-1}\right). \tag{2.2}$$

The mean of this Gaussian distribution corresponds to the cross-sectional maximum likelihood estimate of the factor value plus an adjustment at order $1/n$. The variance shrinks to zero at rate $1/n$ as the class size increases. [8]

Since the conditional distribution of f_t given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}$ is independent of the lagged factor values at order $1/n$, the distribution in (2.2) can be interpreted as an approximation of the filtering distribution for the nonlinear state space model defined by Assumptions A.1-A.3. While the computation of the exact filtering distribution requires in general simulation based methods [see e.g. Duffie et al. (2009) for an application to default models], we exploit the large number n of individual cross-sectional measurements to get a Gaussian approximation at order $1/n$. This approximate Kalman filter shares some common features with the literature on robust Kalman filtering [see e.g. Masreliez (1975)]. However, it differs in several respects. First, in robust filtering the conditional distribution of f_{t+1} given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}$ is assumed to be close to a Gaussian distribution, whereas in our framework it is the conditional distribution of f_t given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}$, which is almost Gaussian [9]. Second, in robust filtering the errors of the analytical approximations are typically unknown [10], while in our approach the Gaussian approximation has been derived theoretically together with its approximation error due to the information contained in the cross-sectional observations. Third, the robust filtering literature mostly focuses on linear measurement and state equations with non-Gaussian innovations [11], while our model fully allows for nonlinearities in both equations. Finally, the approximation in Proposition 2 is not recursive, but in closed form.

Proposition 2 can also be interpreted as an approximation of a posterior distribution in Bayesian statistics [see e.g. Lindley (1980)]. Indeed, let us assume that f_t is an unknown parameter. Then, distribution $g(.|f_{t-1})$ can be interpreted as the prior density, and $E\left[\exp\left(uf_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right]$ characterizes the posterior distribution. Proposition 2 can be seen as an instance of the well-known asymptotic equivalence between Bayesian and frequentist methods in large samples. While in the classical results of Bickel, Yahav (1969) and Ibragimov, Has'minskii (1981) this equivalence is derived directly in terms of the posterior density, we prefer to follow e.g. Lindley (1980) and Tierney, Kadane (1986) and focus on posterior moments. Indeed, the approximate posterior moments of exponential and other nonlinear transformations of f_t are the basis for approximate pricing of derivatives written on the factor and its proxies (see Sections 2.3 and 3). [12] Finally, our Bayesian interpretation considers f_t as the parameter, and is valid for any parametric or nonparametric specification of the factor distribution. It has to be distinguished from Bayesian approaches concerning the mean reversion and volatility parameters of the factor dynamics [see e.g. Duffie et al. (2009), Section 4.4].

The posterior distribution of factor f_t is asymptotically approximated by a well-defined distribution, that is the Gaussian distribution given in (2.2). This is especially important when the approximation concerns the risk-neutral distribution. Indeed, the possibility to interpret the approximation as a probability distribution, that is, with positive density and unit mass, is equivalent to no arbitrage in approximate pricing [13]. From Proposition 2, the conditional expectation of a function $\varphi(f_t)$ of the factor can be approximated at order $1/n$ by computing the integral w.r.t. the Gaussian distribution (2.2).

Corollary 3. *For any integrable function $\varphi\left(f_t\right)$ of f_t, we have:*

$$E\left[\varphi\left(f_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right] = \int \varphi(f_t)\hat{g}_{nt}(f_t)df_t + o(1/n),$$

where $\hat{g}_{n,t}$ is the pdf of the Gaussian distribution (2.2).

The integral in Corollary 3 can be approximated at order $1/n$ in closed form for any smooth function φ.

Corollary 4. *For any twice differentiable function $\varphi\left(f_t\right)$ of f_t, we have:*

$$\begin{aligned} E\left[\varphi\left(f_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right] &= \varphi\left(\hat{f}_{nt}\right) + \frac{1}{n}\frac{d\varphi}{df_t}\left(\hat{f}_{nt}\right)\left[I_{nt}^{-1}\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|\hat{f}_{n,t-1}\right) + \frac{1}{2}I_{nt}^{-2}K_{nt}\right] \\ &\quad + \frac{1}{2n}\frac{d^2\varphi}{df_t^2}\left(\hat{f}_{nt}\right)I_{nt}^{-1} + o(1/n). \end{aligned}$$

Proof. The approximation is derived by expanding function φ at second-order around $\hat{f}_{nt}$:

$$\varphi\left(f_t\right)=\varphi\left(\hat{f}_{nt}\right)+\frac{d\varphi}{df_t}\left(\hat{f}_{nt}\right)\left(f_t-\hat{f}_{nt}\right)+\frac{1}{2}\frac{d^2\varphi}{df_t^2}\left(\hat{f}_{nt}\right)\left(f_t-\hat{f}_{nt}\right)^2+o\left(\left(f_t-\hat{f}_{nt}\right)^2\right),$$

and computing the conditional expectation w.r.t. the density in (2.2). □

Corollary 4 describes how the first- and second-order derivatives of function φ are involved in the expansion of the conditional expectation. This is the analogue of the Ito's formula for a large homogeneous class of assets.

2.2 Examples

The summary statistics $\hat{f}_{n,t}$, $I_{n,t}$ and $K_{n,t}$ for the examples of Section 1.2 are given in Table 1. Let us discuss the associated approximate predictive distributions for Examples (i), (iii) and (v).

i) Linear factor model

For the linear factor model in Example (i), the factor approximation is the cross-sectional average $\hat{f}_{n,t}=\bar{y}_{n,t}$, where $\bar{y}_{n,t}=\frac{1}{n}\sum_{i=1}^n y_{i,t}$. Moreover, $I_{n,t}=1/\sigma^2$ and $K_{n,t}=0$. Let us assume that the latent factor admits a Gaussian autoregressive dynamics $f_t=\mu+\gamma(f_{t-1}-\mu)+\eta\varepsilon_t$, with $\varepsilon_t\sim IIN(0,1)$, unconditional mean μ, and autocorrelation coefficient γ such that $|\gamma|<1$. The approximate predictive distribution of the factor becomes:

$$N\left(\mu+\left(1-\frac{1}{n}\frac{\sigma^2}{\eta^2}\right)\left(\bar{y}_{nt}-\mu\right)+\frac{1}{n}\frac{\sigma^2}{\eta^2}\gamma\left(\bar{y}_{n,t-1}-\mu\right),\frac{\sigma^2}{n}\right).$$

The predictive mean corresponds to the unconditional factor mean μ corrected by a convex combination of $\bar{y}_{nt}-\mu$ and $\gamma\left(\bar{y}_{n,t-1}-\mu\right)$, with weights $1-\frac{1}{n}\frac{\sigma^2}{\eta^2}$ and $\frac{1}{n}\frac{\sigma^2}{\eta^2}$, respectively.

ii) SRF model for default correlation

In the SRF model in Example (iii), the cross-sectional factor approximation is the default frequency $\hat{f}_{n,t}=\bar{y}_{n,t}$. Moreover, $I_{n,t}=1/[\bar{y}_{n,t}(1-\bar{y}_{n,t})]$ and the statistic $K_{n,t}$ does not vanish. Let us assume that the factor F_t admits an autoregressive Gaussian dynamics:

$$F_t=\gamma F_{t-1}+\sqrt{1-\gamma^2}\varepsilon_t, \tag{2.3}$$

where $\varepsilon_t \sim IIN(0,1)$ and $|\gamma| < 1$. The factor is normalized such that the unconditional distribution is standard Gaussian as in the Vasicek (1987) static model. The transition density $g(f_t|f_{t-1})$ of the transformed factor $f_t = \Phi\left(-\frac{\mu + \sigma\sqrt{\rho}F_t}{\sigma\sqrt{1-\rho}}\right)$ is derived by change of variable. We get the approximate predictive distribution:

$$N\left(\hat{f}_{n,t} + \frac{1}{n}\left[\frac{\hat{f}_{n,t}(1-\hat{f}_{n,t})}{\phi\left[\Phi^{-1}(\hat{f}_{n,t})\right]}\left(\Phi^{-1}(\hat{f}_{n,t}) + \frac{\sqrt{1-\rho}}{\sqrt{\rho}}\frac{\hat{F}_{n,t} - \gamma\hat{F}_{n,t-1}}{1-\gamma^2}\right) + 1 - 2\hat{f}_{n,t}\right], \frac{\hat{f}_{n,t}(1-\hat{f}_{n,t})}{n}\right), \tag{2.4}$$

where $\hat{F}_{n,t} = -\frac{1}{\sqrt{\rho}}\left(\frac{\mu}{\sigma} + \sqrt{1-\rho}\Phi^{-1}\left(\hat{f}_{n,t}\right)\right)$ is the cross-sectional approximation of the Gaussian factor. For expository purpose, let us focus on the case of i.i.d. factor $\gamma = 0$, that corresponds to the static Vasicek (1987) model. The predictive mean and variance depend on the available information through the current default frequency $\hat{f}_{n,t} = \bar{y}_{n,t}$ only. Figure 1, middle Panel, displays the predictive mean and the 95% prediction interval as a function of $\hat{f}_{n,t}$ for class size $n = 100$. The parameters are $\rho = 0.10$, $\mu = 0.3501$, $\sigma = 0.10$ (see Section 4 for a discussion of this parameter choice). For comparison, we display in the upper Panel of Figure 1 the predictive mean and 95% prediction intervals for the linear Gaussian single-factor model in Example (i) with $\gamma = 0$ and $n = 100$. The parameters are $\mu = 0.10$, $\eta = 0.055$ and $\sigma = 0.190$ [14]. The predictive distribution features heteroscedasticity for the SRF model, with smaller prediction intervals when $\hat{f}_{n,t}$ is close to either 0 or 1, while the predictive distribution is homoscedastic in the linear model.

iii) Homogenous class of CDS

While in the above examples the factor approximation corresponds to a cross-sectional average of the individual observations, this does not necessarily hold in more complicated factor models. In the homogeneous class of CDS in Example (v), the cross-sectional factor approximation is:

$$\hat{f}_{n,t} = \Psi_\delta^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}[\delta \log y_{i,t} + (1-\delta)\log(1-y_{i,t})]\right), \tag{2.5}$$

where $\Psi_\delta(s) = \delta\Psi(\delta s) + (1-\delta)\Psi((1-\delta)s) - \Psi(s)$ and $\Psi(s) = \frac{d\log\Gamma(s)}{ds}$ is the digamma function (see Appendix 2). For expository purpose, let us consider a static factor model with gamma distribution. More precisely, let us assume that f_t/c is $i.i.d.$ with $\gamma(\nu)$ distribution, where

$c, \nu > 0$ [15]. Then, the predictive distribution is given by:

$$N\left(\hat{f}_{n,t} + \frac{1}{n}\frac{1}{\Psi'_\delta\left(\hat{f}_{n,t}\right)}\left[\frac{\nu+1}{\hat{f}_{n,t}} - 1/c - \frac{1}{2}\frac{\Psi''_\delta\left(\hat{f}_{n,t}\right)}{\Psi'_\delta\left(\hat{f}_{n,t}\right)}\right], \frac{1}{n}\frac{1}{\Psi'_\delta\left(\hat{f}_{n,t}\right)}\right).$$

The predictive mean and the 95% prediction intervals are displayed as a function of $\hat{f}_{n,t}$ in Figure 1, lower Panel, for $n = 100$, $\delta = 0.5$ and $c = \nu = 1$ (exponentially distributed factor). The width of the prediction interval is increasing w.r.t. factor approximation $\hat{f}_{n,t}$.

2.3 Approximate pricing formulas

Let us now consider a European derivative with time-to-maturity 1 and payoff $a\left(y_{1,t+1}\right)$, say. The price of this derivative at date t is:

$$\begin{aligned}
\tilde{\pi}_t(a,1) &= \int \cdots \int a\left(y_{1,t+1}\right) p(y_{1,t+1}, ..., y_{n,t+1}, f_t|\Omega_t) dy_{1,t+1} \cdots dy_{n,t+1} df_t \\
&= \int m(f_t)\left(\int a\left(y_{1,t+1}\right) h\left(y_{1,t+1}|f_{t+1}\right) g(f_{t+1}|f_t) dy_{1,t+h} df_{t+1}\right) g\left(f_t|\Omega_t\right) df_t \\
&= E\left[m(f_t)\alpha(f_t)|\Omega_t\right], \text{ say,} \qquad (2.6)
\end{aligned}$$

by the iterated expectation theorem and equation (1.2). **It is equivalent to price a derivative written on the traded assets with payoff $a\left(y_{1,t+1}\right)$, or to price a virtual derivative written on the unobserved factor value with payoff** [16] $\alpha\left(f_t\right)$. This argument can be extended to any payoff $a\left(y_{1,t+\tau}, ..., y_{n,t+\tau}\right)$ and horizon τ. The variables $y_{i,t}$ do not need to be individual asset prices or returns, but can correspond to individual risk events, as individual default or prepayment for credit, and death or lapse for life insurance.

The approximation theorem in Corollary 3 provides pricing formulas at order $1/n$ for European payoffs written on the factor. Let us first consider the short horizon and denote $\pi_t(\alpha, 0)$ the price at date t of the payoff $\alpha\left(f_t\right)$. We have:

$$\begin{aligned}
\pi_t(\alpha,0) &= E\left[m(f_t)\alpha\left(f_t\right)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right] \\
&= \int m(f_t)\alpha(f_t)\hat{g}_{n,t}(f_t) df_t + o(1/n), \qquad (2.7)
\end{aligned}$$

by Corollary 3. Thus, the derivative price can be computed at order $1/n$ from the observed values $\underline{y_{1,t}}, ..., \underline{y_{n,t}}$ only, whenever the sdf (the risk premia) is given. The approximate pricing formula can be used to define a coherent system of derivative quotations, even if the factor is unobservable at all dates, for instance when no additional derivative is highly traded.

Let us now consider another horizon τ, $\tau \geq 1$, and denote $\pi_t(\alpha, \tau)$ the price at date t of payoff $\alpha\left(f_{t+\tau}\right)$. We have:

$$\pi_t(\alpha, \tau) = E\left[m(f_t)m(f_{t+1}) \cdots m(f_{t+\tau})\alpha\left(f_{t+\tau}\right)|\Omega_t\right].$$

By the iterated expectation theorem, we can first condition on Ω_t, f_t to get:

$$\begin{aligned}\pi_t(\alpha, \tau) &= E\left[E\left[m(f_t)m(f_{t+1}) \cdots m(f_{t+\tau})\alpha\left(f_{t+\tau}\right)|\Omega_t, f_t\right]|\Omega_t\right] \\ &= E\left[m(f_t)\Pi(f_t, \alpha, \tau)|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right] = \pi_t\left[\Pi(\cdot, \alpha, \tau), 0\right],\end{aligned}$$

where $\Pi(f_t, \alpha, \tau) = E\left[m(f_{t+1}) \cdots m(f_{t+\tau})\alpha\left(f_{t+\tau}\right)|\Omega_t, f_t\right]$ is a function of f_t only by the Markov property. Function $\Pi(f_t, \alpha, \tau)$ corresponds to the price at time t of payoff $\alpha\left(f_{t+\tau}\right)$ with time-to-maturity τ, computed by an informed investor who has the larger information set (Ω_t, f_t) available at date t, and uses the sdf $m(f_{t+1})$ to discount risk between t and $t+1$. By Corollary 3 applied to $\varphi(f; \alpha, \tau) = m(f)\Pi(f, \alpha, \tau)$, the derivative price $\pi_t(\alpha, \tau)$ can be approximated at order $1/n$ by using the observed values $\underline{y_{1,t}}, ..., \underline{y_{n,t}}$ only:

$$\pi_t(\alpha, \tau) = \int m(f_t)\Pi(f_t, \alpha, \tau)\hat{g}_{n,t}(f_t)df_t + o(1/n). \tag{2.8}$$

These approximate pricing formulas are applied in the next Section to derivatives written on a factor proxy.

3 Derivatives written on a factor proxy

3.1 Review of the literature on basket derivative pricing

As noted above, the common stochastic factor captures the nondiversifiable component of individual risks, also called systematic risk. It is not surprising to see market solutions to reduce the nondiversifiable risk exposure. Since the underlying factor is not directly observable, it is not possible to propose derivatives written on the factor itself. However, such derivatives can be replaced by basket derivatives written on an observable factor proxy. A typical example is provided by Basket Default Swaps (BDS). These are digital derivatives that pay off at maturity, if the frequency of defaults in a given pool of obligors is larger than a given threshold. Similarly, a tranche of a synthetic Collateralized Debt Obligation (CDO) offers protection when the frequency of default in the pool is in a given range. The frequency of default is a proxy for the systematic credit risk factor.

The literature on basket derivative pricing has mostly focused on static factor models. Hull, White (2004) and Laurent, Gregory (2005) consider copula factor models and introduce semi-analytical and numerical approaches for pricing BDSs and CDOs. Large portfolio approximations are useful to get pricing methods that are less time consuming. Vasicek (1991) introduces a large portfolio approximation in the static SRF model [see Example (iii) of Section 1.2] to approximate the distribution of the default frequency $\hat{f}_{n,t} = \frac{1}{n}\sum_{i=1}^{n} y_{i,t}$ in a pool of obligors. Based on a limit argument for an infinite class (i.e. $n = \infty$), the large portfolio approximation of Vasicek (1991) consists in replacing the distribution of the default frequency $\hat{f}_{n,t}$ by the distribution of the conditional default probability f_t [17]. This corresponds to the infinite granularity assumption of the Basel 2 regulation for credit risk management. For derivative pricing, this large portfolio approximation suggests to price a basket derivative as if it were written on the factor value itself instead of the factor proxy. To adjust for the granularity of the portfolio, the more recent literature considers approximations for large, but finite, cross-sectional dimension n [see e.g. Gordy (2003), (2004) and references therein for granularity adjustments of the Value-at-Risk and expected shortfall of the portfolio loss distribution in a static framework]. The idea is to approximate the conditional distribution of $\hat{f}_{n,t}$ given f_t for large n, and then to integrate out factor f_t. Bastide, Benhamou, Ciuca (2007), El Karoui, Jiao, Kurtz (2008) and El Karoui, Jiao (2009) correct a Gauss (or Poisson) approximation of the distribution of $\hat{f}_{n,t}$ given f_t by using Stein's zero-bias method [18].

While all the above papers consider static factor models only, Lamb, Perraudin, Van Landschoot (2008) extend the SRF model by allowing an autoregressive dynamics for a multivariate factor. The investors' information set at date t includes the factor value f_t only. By following Vasicek (1991), the large portfolio approximation of Lamb, Perraudin, Van Landschoot (2008) is based on the replacement of the conditional distribution of $\hat{f}_{n,t+1}$ given f_t by the conditional distribution of f_{t+1} given f_t. This conditional distribution is used to price synthetic CDO tranches.

Our paper contributes to the literature by considering a rather general factor model specification in a dynamic framework (Assumptions A.1-A.5). The factor value f_t at date t is not included in the investors' information set Ω_t. As a result, the conditional distribution of $\hat{f}_{n,t+1}$ given f_t has to be integrated out w.r.t. the distribution of f_t given Ω_t (see Sections 3.2 and 3.3). This sharply differs from static factor models, where the integration is performed w.r.t. the unconditional distribution of f_t, and from dynamic models where f_t is assumed observable by the investors at date t. In our paper the large portfolio approximation concerns primarily the distribution of f_t given Ω_t (see Proposition 2), and not only the conditional distribution of $\hat{f}_{n,t+1}$ given f_t as in the previous literature.

3.2 Approximate pricing of derivatives written on a default frequency

i) Short-term α-to-default swap

Let us consider a large pool of firms of similar size in a given industrial sector. Let y_{it} denote the indicator for default occurrence, that is $y_{it} = 1$, if firm i defaults at date t, and $= 0$, otherwise. Suppose that the joint distribution of the individual default indicators is given by the SRF model [Example (iii) in Section 1.2], in which $y_{1,t}$, ..., $y_{n,t}$ are i.i.d. with the same Bernoulli distribution $\mathcal{B}(1, f_t)$, conditional on the transformed factor f_t. The Gaussian factor F_t follows the autoregressive dynamics (2.3). A α-to-default swap with maturity $t+1$ pays one Euro, if the fraction of firms in the pool which are in default at $t+1$ is above 100α percent, $\alpha \in (0, 1]$, and pays zero Euro, otherwise. The payoff of this derivative is given by:

$$a\left(y_{1,t+1}, ..., y_{n,t+1}\right) = \mathbb{I}_{\bar{y}_{n,t+1} \geq \alpha},$$

where $\bar{y}_{n,t+1} = \frac{1}{n}\sum_{i=1}^{n} y_{i,t+1}$.

The basket default swap can be interpreted as a derivative written on a factor proxy. Indeed, let $n_t = \sum_{i=1}^{n}(1 - y_{i,t})$ denote the number of firms which are still operating at the end of year t. Without loss of generality, we assume that these firms correspond to indices $i = 1, ..., n_t$. Then, we have $\bar{y}_{n,t+1} = \frac{n_t}{n}\hat{f}_{n,t+1}$, where $\hat{f}_{n,t+1} = \frac{1}{n_t}\sum_{i=1}^{n_t} y_{i,t+1}$ is the cross-sectional ML estimator of default probability f_{t+1} at date $t+1$ in the pool of n_t firms. Estimator $\hat{f}_{n,t+1}$ corresponds to the future default frequency. Thus, the payoff of the α-to-default swap can be written as:

$$a\left(y_{1,t+1}, ..., y_{n,t+1}\right) = \mathbb{1}_{\hat{f}_{n,t+1} \geq \alpha_t},$$

where $\alpha_t = \frac{n}{n_t}\alpha$ is known at date t.

Let us assume for expository purpose a constant sdf and a zero risk-free rate, that is, equal risk-neutral and historical distributions. The price at time t of the α-to-default swap for maturity $t+1$ is given by:

$$p_{n,t}(\alpha, 1) = E\left[\mathbb{1}_{\hat{f}_{n,t+1} \geq \alpha_t} | \Omega_t\right]. \tag{3.1}$$

This price can be computed from the prices of exponential derivatives written on $\hat{f}_{n,t+1}$ by using the Fourier Transform Inversion formula [Duffie, Pan, Singleton (2000)]. More precisely, the price at date t of the derivative with exponential payoff $\exp\left(u\hat{f}_{n,t+1}\right)$ at $t+1$ is given by:

$$\begin{aligned}
\tilde{\pi}_{n,t}(u, 1) &= E\left[\exp\left(u\hat{f}_{n,t+1}\right) | \Omega_t\right] \\
&= E_t\left[E_t\left[\exp\left(u\hat{f}_{n,t+1}\right) | f_{t+1}\right]\right] \\
&= E_t\left[\left\{1 + \left(\exp\left(u/n_t\right) - 1\right) f_{t+1}\right\}^{n_t}\right],
\end{aligned} \tag{3.2}$$

where $E_t[.]$ denotes the conditional expectation given the investors' information $\Omega_t = \left(\underline{y_{1,t}}, \cdots, \underline{y_{n,t}}, \underline{f_{t-1}}\right)$. By the iterated expectation theorem, we get:

$$\tilde{\pi}_{n,t}(u, 1) = E\left[\varphi(f_t; u) | \Omega_t\right], \tag{3.3}$$

where $\varphi(f_t; u) = E\left[\{1 + (\exp(u/n_t) - 1) f_{t+1}\}^{n_t} \mid f_t\right]$. The density of f_t given Ω_t is computed by formula (1.3) applied to the pool of n_{t-1} firms operating in period $(t-1, t)$:

$$g\left(f_t|\Omega_t\right) = \frac{g(f_t|f_{t-1})\left(f_t\right)^{n_{t-1}-n_t}\left(1-f_t\right)^{n_t}}{\displaystyle\int g(f_t|f_{t-1})\left(f_t\right)^{n_{t-1}-n_t}\left(1-f_t\right)^{n_t} df_t}. \tag{3.4}$$

Thus, the true derivative price $\tilde{\pi}_{n,t}(u,1)$ depends on the investor information Ω_t through f_{t-1}, n_t and n_{t-1}. Then, from the Fourier Transform Inversion formula in Proposition 2 of Duffie, Pan, Singleton (2000), which applies to purely imaginary argument $u = iv$, we get [19]:

$$p_{n,t}(\alpha, 1) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{Im\left[\tilde{\pi}_{n,t}(iv,1)\exp\left(-iv\alpha_t\right)\right]}{v} dv, \tag{3.5}$$

where Im denotes the imaginary component of a complex number.

The large portfolio approximation of the α-to-default swap price is derived by the approximation of the exponential derivative prices. By Corollary 3 and equation (3.3), for $n_{t-1} \to \infty$ we get:

$$\tilde{\pi}_{n,t}(u,1) = \int \varphi(f_t; u)\hat{g}_{nt}(f_t) df_t + o(1/n), \tag{3.6}$$

where $\hat{g}_{nt}$ is the pdf of the normal distribution in (2.4) with $n = n_{t-1}$ and $\hat{f}_{n,t} = \frac{1}{n_{t-1}}\sum_{i=1}^{n_{t-1}} y_{i,t}$. Since $\hat{f}_{nt} = (n_{t-1} - n_t)/n_{t-1}$, the approximate derivative price depends on the past default history through the counts n_{t-2}, n_{t-1} and n_t of firms operating at the end of year $t-2$, $t-1$ and t, respectively. An equivalent summary of past default history is n_{t-1}, n_t, $\hat{f}_{n,t-1}$ [20] . Then, the approximation of the α-to-default swap price is obtained by the Fourier Transform Inversion formula (3.5) applied with approximation (3.6) evaluated at purely imaginary argument.

ii) α-to-default swap at longer horizon

Let us now consider a α-to-default swap with another time-to-maturity τ. The payoff at $t+\tau$ is given by:

$$a\left(y_{1,t+\tau}, ..., y_{n,t+\tau}\right) = \mathbb{I}_{\bar{y}_{n,t+\tau} \geq \alpha}.$$

The approximation of the derivative price can be derived along the same lines as for short horizon $\tau = 1$. Indeed, conditional on a factor path, the default probability at horizon τ for a firm which is operating at t is:

$$P\left[y_{i,t+\tau}=1|y_{i,t}=0,(f_t)\right]=1-(1-f_{t+1})\cdots(1-f_{t+\tau})=:\lambda_{t+\tau}.$$

Thus, conditional on a factor path, the default indicators $y_{i,t+\tau}$, $i=1,...,n_t$, of the firms operating at t are i.i.d. with Bernoulli distribution $\mathcal{B}\left(1,\lambda_{t+\tau}\right)$. The same approach as above can be applied, with $\varphi(f_t;u)=E\left[\left\{1+\left(\exp\left(u/n_t\right)-1\right)\lambda_{t+\tau}\right\}^{n_t}\mid f_t\right]$.

iii) Other examples of derivatives written on a default frequency

The approximate pricing approach introduced for α-to-default swaps can be extended to other derivatives written on a default frequency. For instance, a synthetic CDO tranche is a derivative that offers protection against the portfolio losses in a specific range. Let us consider a portfolio with homogeneous nominals and zero recovery rates. The percentage portfolio loss at horizon τ is $\hat{f}_{n,t+\tau}$, and the payoff is $\left(\hat{f}_{n,t+\tau}-\alpha_1\right)^+-\left(\hat{f}_{n,t+\tau}-\alpha_2\right)^+$, where $\alpha_1<\alpha_2$ are called attachment (resp. detachment) points. An actively traded contract of this type is written for instance on the iTraxx Europe index, which is an index based on the CDS spreads of 125 companies. Finally, derivatives written on an observable factor proxy have been proposed in the insurance industry. A typical example is the longevity bond. Such a bond has a contractual maturity (e.g. 25-year) and pays regularly a coupon proportional to the current frequency of surviving people in a contractual cohort, such as a national population with given age at the date of bond issuing. These observed frequencies are proxies of the successive values of a longevity factor and correspond to the cross-sectional maximum likelihood estimates of these factor values. Approximate pricing formulas for longevity bonds can be derived by considering for instance a single-factor model similar to the SRF model in Sections 3.2 i), ii), where the common factor is assumed nonstationary to capture a stochastic trend in mortality risk.

3.3 Granularity adjustment for derivatives written on a factor proxy

While in the SRF model for default the factor approximation is a cross-sectional average, in a more complex factor model the factor approximation does not admit in general the interpretation of a cross-sectional average. For instance, for a homogeneous class of CDS such as Example (v) in

Section 1.2, the cross-sectional factor approximation is a nonlinear transformation of the quantity $\frac{1}{n}\sum_{i=1}^{n}[\delta \log y_{i,t} + (1-\delta)\log(1-y_{i,t})]$, see equation (2.5). In this case, the relevant proxy of the systematic risk factor may be very different from a cross-sectional average of the CDS prices.

In this Subsection we derive approximate pricing formulas for the general framework of a homogeneous class of assets satisfying Assumptions A.1-A.7 and derivatives written on the factor proxy $\hat{f}_{n,t+\tau} = \arg\max_{f_{t+\tau}} \sum_{i=1}^{n} \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)$, that is the cross-sectional ML estimator of the factor value $f_{t+\tau}$ at time-to-maturity τ. We focus on approximate pricing of derivatives with exponential payoff $\exp\left(u\hat{f}_{n,t+\tau}\right)$. These derivatives are the basis for approximate pricing of more general payoffs by means of the Fourier Transform Inversion formula (see Section 3.2).

Proposition 5. *The true price $\tilde{\pi}_{n,t}(u,\tau)$ at time t of the derivative with payoff $\exp\left(u\hat{f}_{n,t+\tau}\right)$ at $t+\tau$ is such that:*

$$\tilde{\pi}_{n,t}(u,\tau) = E\left[\varphi_n\left(f_t; u, \tau\right)|\Omega_t\right] + o(1/n), \tag{3.7}$$

where:

$$\varphi_n\left(f_t; u,\tau\right) = E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})\exp\left(uf_{t+\tau} - \frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau} + \frac{u^2}{2n}I_{t+\tau}^{-1}\right)|f_t\right], \tag{3.8}$$

with $I_{t+\tau} = E\left[-\frac{\partial^2 \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f^2}|f_{t+\tau}\right]$ and:

$$\beta_{t+\tau} = Cov\left(\frac{\partial \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f}, \frac{\partial^2 \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f^2} + \left(\frac{\partial \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f}\right)^2 |f_{t+\tau}\right).$$

Proof. See Appendix 3. ☐

From equation (2.6), we know that the derivative with payoff $\exp\left(u\hat{f}_{n,t+\tau}\right)$ is equivalent to a virtual derivative with payoff written on the factor value. Proposition 5 shows that, at order $1/n$, this payoff does not involve large dimensional integrals w.r.t. the future values of variables $y_{i,t}$, $i = 1, ..., n$, but only an expectation w.r.t. the factor path.

The quantity $\exp\left(uf_{t+\tau} - \frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau} + \frac{u^2}{2n}I_{t+\tau}^{-1}\right)$ in equation (3.8), u varying, is the Laplace transform of $\hat{f}_{n,t+\tau}$ conditional on $f_{t+\tau}$ at order $1/n$. It corresponds to the Gaussian distribution $N\left(f_{t+\tau} - \frac{1}{2n}I_{t+\tau}^{-2}\beta_{t+\tau}, \frac{1}{n}I_{t+\tau}^{-1}\right)$. The terms of order $1/n$ capture the effect of finite cohort size.

More precisely, the price of the virtual derivative written on the factor value itself is:

$$\pi_{\infty,t}(u,\tau) = E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})\exp\left(uf_{t+\tau}\right)|\Omega_t\right].$$

The difference between the derivative prices $\tilde{\pi}_{n,t}(u,\tau)$ and $\pi_{\infty,t}(u,\tau)$ is the granularity adjustment.

The large portfolio approximation theorems of Section 2 can be used to approximate the derivative price $\tilde{\pi}_{n,t}(u,\tau)$. By Corollary 3 we get the next result.

Corollary 6. *An approximation at order* $1/n$ *of derivative price* $\tilde{\pi}_{n,t}(u,\tau)$ *is:*

$$\tilde{\pi}_{n,t}(u,\tau) \quad = \quad \int \varphi_n(f_t;u,\tau)\hat{g}_{nt}(f_t)df_t + o(1/n), \tag{3.9}$$

where $\varphi_n\left(f_t;u,\tau\right)$ *is given in (3.8) and* $\hat{g}_{nt}$ *is the pdf of the Gaussian distribution (2.2).*

The approximate derivative price in Corollary 6 depends only on the observed $y_{i,t}$'s.

Let us illustrate the use of the pricing formulas in two examples of Section 1.2 (see Table 1 for the statistics β_t and I_t in the other examples).

i) SRF model for default correlation: We have $I_{t+1} = 1/[f_{t+1}(1-f_{t+1})]$ and $\beta_{t+1} = 0$. Then, from Proposition 5 the true price of the exponential derivative with time-to-maturity 1 is $\tilde{\pi}_{n,t}(u,1) = E\left[\varphi_n(f_t;u)|\Omega_t\right]$ at order $1/n$, where $\varphi_n(f_t;u) = m(f_t)E\left[\exp\left(uf_{t+1}+\frac{u^2}{2n}f_{t+1}(1-f_{t+1})\right)|f_t\right]$ and $n = n_t$ [21]. The ratio between $\tilde{\pi}_{n,t}(u,1)$ and the price $\pi_{\infty,t}(u,1) = E\left[m(f_t)\exp(uf_{t+1})|\Omega_t\right]$ of the virtual derivative written on the factor value is:

$$\frac{\tilde{\pi}_{n,t}(u,1)}{\pi_{\infty,t}(u,1)} \quad = \quad E_t^{P_u}\left[\exp\left(\frac{u^2}{2n}f_{t+1}\left(1-f_{t+1}\right)\right)\right] + o(1/n),$$

where P_u is a modified probability with density $m(f_t)\exp\left(uf_{t+1}\right)/E_t\left[m(f_t)\exp\left(uf_{t+1}\right)\right]$. In particular, the price $\tilde{\pi}_{n,t}(u,1)$ is always larger than $\pi_{\infty,t}(u,1)$ for large n, and decreases with the size n of the underlying cohort. Since $V_t\left[\hat{f}_{n,t+1}|f_{t+1}\right] = \frac{1}{n}f_{t+1}\left(1-f_{t+1}\right)$, the ratio $\tilde{\pi}_{n,t}(u,1)/\pi_{\infty,t}(u,1)$ represents the price of the aggregate idiosyncratic risk, which does not vanish when the cohort has a finite size. We get an approximate derivative price from equation (3.9), where $\hat{g}_{n,t}$ is the Gaussian distribution (2.4) with $n = n_{t-1}$.

ii) Homogenous class of CDS: Derivatives based on a class of CDS have been introduced on the market and are generally written on an average of CDS spreads. In a single-factor model, in which the systematic factor measures the concentration of risk, such an average does not represent the appropriate cross-sectional factor proxy. Proposition 5 and Corollary 6 suggest new derivative designs, where the derivative payoff depends on a cross-sectional maximum likelihood approximation of the factor. For the homogeneous class of CDS in Example (v) in Section 1.2, we have $I_{t+1} = \Psi'_\delta(f_{t+1})$ and $\beta_{t+1} = \Psi''_\delta(f_{t+1})$ (see Appendix 2). The price of exponential derivatives written on the factor proxy $\hat{f}_{n,t+1}$ are computed as conditional expectation of the discounted payoff:

$$\varphi_n(f_t; u) = m(f_t)E\left[\exp\left(uf_{t+1} - \frac{u}{2n}\frac{\Psi''_\delta(f_{t+1})}{[\Psi'_\delta(f_{t+1})]^2} + \frac{u^2}{2n}\frac{1}{\Psi'_\delta(f_{t+1})}\right)|f_t\right].$$

The (theoretical) granularity adjustment is:

$$\frac{\tilde{\pi}_{n,t}(u,1)}{\pi_{\infty,t}(u,1)} = E_t^{P_u}\left[\exp\left(-\frac{u}{2n}\frac{\Psi''_\delta(f_{t+1})}{\Psi'_\delta(f_{t+1})^2} + \frac{u^2}{2n}\frac{1}{\Psi'_\delta(f_{t+1})}\right)\right] + o(1/n).$$

It involves mean and variance adjustments, associated with the terms in u and u^2, respectively. The presence of a mean adjustment may imply a ratio of prices smaller than 1, that is a negative granularity adjustment. Indeed, for small u we have $\frac{\tilde{\pi}_{n,t}(u,1)}{\pi_{\infty,t}(u,1)} \simeq 1 - \frac{u}{2n}E_t\left[\frac{\Psi''_\delta(f_{t+1})}{\Psi'_\delta(f_{t+1})^2}\right]$. Since $E_t\left[\frac{\Psi''_\delta(f_{t+1})}{\Psi'_\delta(f_{t+1})^2}\right] \neq 0$, the ratio of prices is smaller than 1 for some derivatives. Therefore, the price $\tilde{\pi}_{n,t}(u,1)$ of the derivative with payoff $\exp(u\hat{f}_{n,t+1})$ is not necessarily a decreasing function of n, or equivalently the price $\pi_{\infty,t}(u,1)$ is not a lower bound for price $\tilde{\pi}_{n,t}(u,1)$. Even if the underlying portfolio gets more diversified when size n increases, the above unexpected effect can be explained by the asymptotic bias in the proxy $\hat{f}_{n,t+1}$, whenever $\beta_{t+1} \neq 0$. This is a rather common feature of nonlinear factor models.

4 Numerical illustration to basket default swap

In this section we provide a numerical illustration to approximate pricing of α-to-default swap derivatives within a dynamic SRF model (see Sections 2.2 and 3.2). Compared to the standard pricing methods based on static factor models, our approach accounts for all individual default histories. We focus on the pricing error due to the large portfolio approximation of the factor predictive distribution. We do not analyze the granularity adjustment which measures the difference between the true prices and the misspecified pricing with $n = \infty$ proposed in Basel 2.

4.1 The SRF model

The time period corresponds to one year. In SRF model (1.4), the asset volatility is $\sigma = 0.20$, and parameter μ is set equal to 0.3501 in order to match an historical (unconditional) default probability of 4% [22] . The asset correlation parameter ρ can be calibrated in order to match relevant values of (unconditional) default correlation [23] . In the empirical literature, different orders of magnitude for estimated default correlations have been proposed, according to the country, firm size and characteristics used to group the firms in homogeneous classes. For instance, values of about 1% have been found when large US firms are grouped into industrial sectors [De Servigny, Renault (2002), Feng, Gouriéroux, Jasiak (2008)], while the estimated default correlations are about 0.1% for small and medium size French firms classified according to both industrial sector and rating [Gagliardini, Gouriéroux (2005)]. These values of default correlation imply an asset correlation ρ below 0.10, much smaller than the value of about 0.30 given by the formula proposed by the Basle Committee for a default probability of 4%. To cover the range obtained with these different approaches, we consider three values of asset correlation, that are $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$, respectively. Finally, the common factor F_t is a Gaussian autoregressive process as in (2.3), where the autocorrelation coefficient is $\gamma = 0.5$ and the innovations η_t are i.i.d. standard Gaussian variables. For expository purpose, the sdf is assumed constant with zero risk-free rate.

4.2 Factor distributions and derivative prices

Let us first illustrate the patterns of the factor distribution and α-to-default swap price for a specific past default history. The numbers of operating firms at the end of years $t-1$ and t are $n_{t-1} = 1000$ and $n_t = 960$, respectively, which imply $\hat{f}_{nt} = 0.04$. We consider three different values of n_{t-2}, which correspond to $\hat{f}_{n,t-1} = 0.04$, $\hat{f}_{n,t-1} = 0.0025$, and $\hat{f}_{n,t-1} = 0.125$, respectively.

Figure 2 displays the conditional distribution of factor f_t given f_{t-1}, for different values of f_{t-1} and asset correlation $\rho = 0.10$. The conditioning values of f_{t-1} are given in terms of their corresponding Gaussian factor F_{t-1}; they are $F_{t-1} = 0$, $F_{t-1} = 2$, $F_{t-1} = -2$, respectively. Figure 3 displays the conditional distribution of factor f_t given the investors' information f_{t-1}, n_{t-1}, and n_t, for different values of f_{t-1} and asset correlation $\rho = 0.10$. The conditioning values of the lagged factor are $F_{t-1} = 0$, $F_{t-1} = 2$ and $F_{t-1} = -2$, respectively. The density is obtained by formula (3.4), where the integral in the denominator is computed numerically. Figure 4 displays the approximate distribution of factor f_t given the past default history n_t, n_{t-1}, $\hat{f}_{n,t-1}$, for different values of $\hat{f}_{n,t-1}$. Asset correlation is $\rho = 0.10$. The approximation is obtained by formula (2.4). Whereas the sole knowledge of the lagged factor value results in skewed distributions with very different patterns (Figure 2), the observed default frequencies are very informative. When this additional information is introduced, the predictive distributions given the investors' information are close to Gaussian distributions, peaked at a value near $\hat{f}_{n,t} = 0.04$ and much less sensitive to the lagged factor value (see Figure 3). Similarly, the approximate predictive distributions given the default history are not very sensitive to the estimated lagged factor value (see Figure 4), and close to the predictive distributions given the investors' information displayed in Figure 3. These findings are consistent with Proposition 2.

The true price of α-to-default swap at time-to-maturity 1 year with payoff $\mathbb{1}_{\hat{f}_{n,t+1} \geq \alpha n/n_t}$ [see Section 3.2 i)] is displayed in Figure 5 as a function of α, for three different values of asset correlation $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$, respectively. The initial size of the pool is $n = 1000$, and the current size is $n_t = 960$. The lagged Gaussian factor value is $F_{t-1} = 0$. The true price is computed with the Fourier Transform Inversion formula (3.5), where the exponential derivative prices are obtained from (3.2) by Monte-Carlo simulation with the acceptance-rejection algorithm based on $10,000$ draws [see e.g. Robert, Casella (2004) and Appendix 4 i) for a discussion of this algorithm]. The pattern of the true price as a function of α corresponds to the (risk-neutral) survivor function of the future factor proxy $\hat{f}_{n,t+1}$ [see equation (3.1)]. For small values of ρ, this function is close to a Gaussian survivor function because of the Central Limit Theorem [24]. Moreover, the derivative prices for large values of α, above 10% say, are very small, since they correspond to rare joint default events. As expected, default correlation is a key parameter to measure the quality of a basket derivative. The true price of the α-to-default swap decreases in ρ (resp. increases in ρ) for small (resp. large) values of α. This is due to the positive effect of ρ on the variance of the (risk-neutral) distribution of $\hat{f}_{n,t+1}$. Figure 6 displays the true price of the α-to-default swap at time-to-maturity 1 year as a function of α, for three different values of the lagged factor $F_{t-1} = 0$, $F_{t-1} = -2$, and $F_{t-1} = 2$. Asset correlation is $\rho = 0.10$. Since the observed current and past default frequencies are included in the investor's information set, the derivative price is not very sensitive to the lagged factor value.

Figure 7 displays the approximate price of the α-to-default swap at time-to-maturity 1-year as a function of α, for three different values of asset correlation $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$. The past default history is such that $\hat{f}_{n,t-1} = 0.04$. The approximation is obtained by Fourier Transform Inversion formula (3.5) and approximation (3.6) for exponential derivatives [see also Appendix 4 ii) for the implementation]. The approximate prices with $\hat{f}_{n,t-1} = 0.04$ are close to the true prices with $F_{t-1} = 0$ (see Figure 5). Finally, Figure 8 displays the approximate price of the α-to-default swap at time-to-maturity 1 year as a function of α, for asset correlation $\rho = 0.10$ and three different values of $\hat{f}_{n,t-1}$. The approximate price is not very sensitive to the estimate of the lagged factor value.

4.3 Monte-Carlo

Let us now investigate the size of the pricing errors implied by the approximation formula. Let us consider a α-to-default contract issued at date $t = 0$ on a pool of $n = n_0$ obligors and maturing at date $t = 3$. We consider two different values for the initial size of the pool, which are $n = 100$ and $n = 1000$. The price is computed at date $t = 2$, for default frequencies $\alpha = 2.5\%$, $\alpha = 5\%$, $\alpha = 10\%$ and $\alpha = 12.5\%$. To compare the true and approximate prices, we perform a Monte-Carlo experiment to simulate the factor path f_t, $t = 1, 2$, and the default history summarized by n_t, $t = 1, 2$. The price approximation is not valid when $\hat{f}_{n,t}$ is equal to either 0 or 1, where the approximate factor density is degenerate with zero variance. Thus, we disregard the realizations with either $\hat{f}_{n,t} = 0$, or $\hat{f}_{n,t} = 1$, which amounts to simulate the process conditional on the event $\hat{f}_{n,t} \in (0, 1)$. For each Monte-Carlo replication, we compute i) the true prices at $t = 2$ using n_t, n_{t-1}, f_{t-1}, and ii) the approximate prices using $n_t, n_{t-1}, \hat{f}_{n,t-1}$. We perform 2500 replications. The computation of the theoretical and approximated derivative prices in one replication takes about 4 seconds and 40 seconds, respectively, on a standard computer.

Table 2 reports the mean, the median, the two quartiles and the upper and lower 5% quantiles of the distribution of the relative pricing errors, as well as the median of the absolute value of the relative pricing errors. The three panels refer to asset correlation $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$, respectively. The relative pricing error is defined as the difference between approximated price and true price, divided by the true price. The approximate prices are obtained by the method discussed in Section 4.2. The bias is rather small, and increases for small size of the class and large asset correlation [25] (see the Panel with $\rho = 0.30$ and $n = 100$). For default frequencies $\alpha = 2.5\%$, $\alpha = 5\%$, asset correlations $\rho = 0.10$, $\rho = 0.30$ and class size $n = 1000$, the median absolute value of the relative pricing errors in percentage is below 3%. Relative pricing errors in percentage are less than 10% in absolute value with probability at least 0.90. At the contrary, the median absolute value of the relative pricing errors in percentage may be larger than 5% for either frequencies $\alpha = 10\%$, $\alpha = 12.5\%$, or for class size $n = 100$, or for asset correlation $\rho = 0.01$. To explain these findings, note that frequencies $\alpha = 10\%$ and $\alpha = 12.5\%$ correspond to rather extreme joint default events compared to the historical default probability of 4%. The associated true derivative prices are often very close to zero. In practice, since prices are displayed in discrete ticks, these small values correspond to zero prices, yielding infinite relative approximation errors. To avoid this problem, the statistics in Table 2 are computed using only the Monte-Carlo values with true prices larger than .005. Still, for true prices close to this lower bound, the relative approximation errors are quite large and explain the large values of the quantiles displayed in Table 2 for frequencies $\alpha = 10\%, 12.5\%$. For asset correlation $\rho = 0.01$, only the statistics for frequencies $\alpha = 2.5\%$, $\alpha = 5\%$ and pool size $n = 1000$ are displayed. Indeed, for $\alpha = 10\%$ and $\alpha = 12.5\%$ the derivative prices are always below .005 (see Figure 5). Moreover for $\rho = 0.01$ the relative pricing errors are rather large, and often above 100% for pool size $n = 100$ (not displayed). This is because for $\rho = 0$ the factor f_t does not impact the individual default indicators, and then the approximation theorem breaks down.

5 Concluding remarks

A large variety of basket derivatives with nonlinear payoffs have been introduced in the market to capture the nonlinear dynamic features of common latent risk factors influencing a homogeneous pool of individual contracts. The associated securitization usually concerns either individual loans, or insurance contracts. In our paper, we have derived approximated derivative pricing formulas for large underlying pools. These formulas are semi-analytic. They can be applied to factors with different interpretations, such as mean, volatility, default correlation, or concentration, and to any factor dynamics. Moreover, the approach considers coherently the historical and risk-neutral dynamics.

The large-portfolio pricing formulas do not involve the unobservable factor values, which are replaced by well-chosen summaries constructed from the observable asset returns. Thus, these approximate pricing formulas can be used even when no price of additional highly traded derivatives is observed, as it is the case at the emergence of a new derivative market. It is known that some major financial risks are due to systematic factors. The approximate pricing formulas can be used to hedge systematic risk by introducing derivatives written on a cross-sectional factor proxy compatible with the factor interpretation.

The present paper focuses on derivative pricing, and assumes the historical model parameters and the risk premia known. In practice, the distribution of the observable variable y_{it} given the factor value f_t, and the transition of the factor f_t, involve unknown parameters. The estimation of these historical parameters is the subject of a companion paper [see Gagliardini, Gouriéroux (2008)]. Similarly, the sdf specification involves unknown risk premium parameters. The choice of these sdf parameters is still an open question. At the emergence of a new derivative market, risk premia are selected by the first firm proposing the products on a monopolistic and risk attitude basis. In a later stage of market development, risk premia will be updated by using the incoming information on the market prices of highly traded derivatives.

Finally, the methodology of the present paper can be extended to more complicated dynamic models: first, to models featuring an idiosyncratic dynamics, that is, an additional effect of lagged $y_{i,t-1}$ on $y_{i,t}$ [see Gagliardini, Gouriéroux, Monfort (2010) for the extension of the approximation formula]; second, to models including both observable and unobservable common factors [see Duffie et al. (2009) for an example of such a model]; third, to models with observed heterogeneity to account for the so-called concentration risk.

Figure Legends

Figure 1: In each Panel, the solid line is the predictive mean for the factor value, and the dashed lines give the 95% prediction interval with class size $n = 100$. We consider static factor models. The available information corresponds to the factor approximation $\hat{f}_{n,t}$. For the single-factor linear model (upper Panel), we assume $f_t \sim IIN(\mu, \eta^2)$, where $\mu = 0.10$ and $\eta = 0.0155$. For the SRF model (middle Panel), we have $F_t \sim IIN(0, 1)$ as in Vasicek (1987). For the homogenous class of CDS (lower Panel), we assume $f_t \sim i.i.d.\gamma(1)$.

Figure 2: The Figure plots the conditional distribution of factor f_t given f_{t-1}, for different values of f_{t-1}. The conditioning values of f_{t-1} are given in terms of their corresponding Gaussian factor values F_{t-1}; they are $F_{t-1} = 0$, $F_{t-1} = 2$, $F_{t-1} = -2$, respectively. Asset correlation is $\rho = 0.10$.

Figure 3: The Figure plots the conditional distribution of factor f_t given f_{t-1}, $n_{t-1} = 1000$, and $n_t = 960$, for different values of f_{t-1}. The conditioning values of f_{t-1} are given in terms of their corresponding Gaussian factor values F_{t-1}; they are $F_{t-1} = 0$, $F_{t-1} = 2$, and $F_{t-1} = -2$ respectively. Asset correlation is $\rho = 0.10$.

Figure 4: The Figure plots the approximate conditional distribution of f_t given past default history $n_{t-1} = 1000$, $n_t = 960$, $\hat{f}_{n,t-1}$, for different values of $\hat{f}_{n,t-1}$. Asset correlation is $\rho = 0.10$.

Figure 5: The Figure plots the price of α-to-default swap at time-to-maturity 1 year as a function of α, for three different values of asset correlation $\rho = 0.01$ (dotted line), $\rho = 0.10$ (solid line) and $\rho = 0.30$ (dashed line). The lagged Gaussian factor value is $F_{t-1} = 0$.

Figure 6: The Figure plots the price of α-to-default swap at time-to-maturity 1 year as a function of α, for three different values of the lagged factor $F_{t-1} = 0$, $F_{t-1} = -2$, and $F_{t-1} = 2$. Asset correlation is $\rho = 0.10$.

Figure 7: The Figure plots the approximate price of α-to-default swap at time-to-maturity 1 year as a function of α, for three different values of asset correlation $\rho = 0.01$ (dotted line), $\rho = 0.10$ (solid line) and $\rho = 0.30$ (dashed line). The past default history is such that $\hat{f}_{n,t-1} = 0.04$.

Figure 8: The Figure plots the approximate price of α-to-default swap at time-to-maturity 1 year as a function of α, for different values of $\hat{f}_{n,t-1}$. Asset correlation $\rho = 0.10$.

References

[1] Bank for International Settlements. 2010. “OTC Derivative Market Activity in the Second Half of 2009.” Monetary and Economic Department, May.

[2] Basel Committee on Banking Supervision. 2001. ”The New Basel Capital Accord.” Consultative Document of the Bank for International Settlements.

[3] Bastide, D., E. Benhamou, and M. Ciuca. 2007. ”A Comparative Analysis of Basket Default Swaps Pricing Using the Stein Method.” Working Paper, Pricing Partners.

[4] Bates, D. 2009. “U.S. Stock Market Crash Risk, 1926-2006.” Working Paper, University of Iowa.

[5] Bickel, P., and J. Yahav. 1969. “Some Contributions to the Asymptotic Theory of Bayes Solutions.” *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 11: 257-276.

[6] Cairns, A., D. Blake, and K. Dowd. 2006. ”A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration.” *Journal of Risk and Insurance* 73: 687-718.

[7] CEA. 2007. “Solvency 2: Understanding the Process.” CEA Discussion Paper, February.

[8] Cipra, T., and A. Rubio. 1991. ”Kalman Filter with a Nonlinear Non-Gaussian Observation Relation.” *Trabajos de Estadistica* 6: 111-119.

[9] Dahl, M., and T. Moller. 2004. ”Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk.” *Insurance: Mathematics and Economics* 39: 193-217.

[10] Darolles, S., C. Gouriéroux, and J. Jasiak. 2006. ”Structural Laplace Transform and Compound Autoregressive Models.” *Journal of Time Series Analysis* 27: 477-503.

[11] de Finetti, B. 1931. ”Funzione Caratteristica di un Fenomeno Aleatorio.” *Atti della R. Accademia dei Lincei* 6, Memorie, Classe di Scienze Fisiche, Matematiche e Naturali 4: 251-299.

[12] De Servigny, A., and O. Renault. 2002. ”Default Correlation: Empirical Evidence.” S&P, October.

[13] Dembo, A., J.-D. Deuschel, and D. Duffie. 2004. ”Large Portfolio Losses.” *Finance and Stochastics* 8: 3-16.

[14] Duffie, D., A. Eckner, G. Horel, and L. Saita. 2009. ”Frailty Correlated Default.” *Journal of Finance* 64: 2089-2123.

[15] Duffie, D., D. Filipovic, and W. Schachermayer. 2003. ”Affine Processes and Applications to Finance.” *Annals of Applied Probability* 13: 984-1053.

[16] Duffie, D., J. Pan, and K. Singleton. 2000. ”Transform Analysis and Asset Pricing for Affine Jump Diffusions.” *Econometrica* 68: 1343-1376.

[17] Duffie, D., and K. Singleton. 2003. *Credit Risk: Pricing, Measurement and Management.* Princeton, NJ: Princeton University Press.

[18] El Karoui, N., and Y. Jiao. 2009. “Stein’s Method and Zero Bias Transformation for CDO Tranche Pricing.” *Finance and Stochastic* 13: 151-180.

[19] El Karoui, N., Y. Jiao, and D. Kurtz. 2008. ”Gauss and Poisson Approximation: Application to CDOs Tranche Pricing.“ *Journal of Computational Finance* 12: 2.

[20] Feng, D., C. Gouriéroux, and J. Jasiak. 2008. ”The Ordered Qualitative Model for Credit Rating Transitions.” *Journal of Empirical Finance* 15:111-130.

[21] Frey, R., and A. Mc Neil. 2003. ”Dependent Defaults in Models of Portfolio Credit Risk.” *Journal of Risk* 6: 59-62.

[22] Gagliardini, P., and C. Gouriéroux. 2005. ”Stochastic Migration Models with Application to Corporate Risk.” *Journal of Financial Econometrics* 3: 188-226.

[23] Gagliardini, P., and C. Gouriéroux. 2008. ”Efficiency in Large Dynamic Panel Models with Common Factor.” Working Paper, CREST and Swiss Finance Institute.

[24] Gagliardini, P., C. Gouriéroux, and A. Monfort. 2010. “Microinformation, Nonlinear Filtering and Granularity.” Working Paper, CREST and Swiss Finance Institute.

[25] Gordy, M. 2002. "Saddlepoint Approximation of CreditRisk+." *Journal of Banking and Finance* 26: 1335-1353.

[26] Gordy, M. 2003. "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules." *Journal of Financial Intermediation* 12: 199-232.

[27] Gordy, M. 2004. "Granularity Adjustment in Portfolio Risk Measurement." In G. Szegö (ed.), *Risk Measures for the 21th Century*, Wiley.

[28] Gouriéroux, C. 2008. "CaR and Affine Processes." In R. Mariano and Y. Tse (eds.), *Econometric Forecasting and High Frequency Data Analysis*, Lecture Notes Series, Institute of Mathematical Science Singapore.

[29] Gouriéroux, C., and J. Jasiak. 2006. "Autoregressive Gamma Process."*Journal of Forecasting* 25: 129-152.

[30] Gouriéroux, C., and A. Monfort. 1995. *Statistics and Econometric Models.* Cambridge: Cambridge University Press.

[31] Gouriéroux, C., and A. Monfort. 2008. "Quadratic Stochastic Intensity and Prospective Mortality Tables." *Insurance: Mathematics and Economics* 43: 174-184.

[32] Gouriéroux, C., A. Monfort, and V. Polimenis. 2006. "Affine Term Structure Models for Credit Risk." *Journal of Financial Econometrics* 4: 31-52.

[33] Gouriéroux, C., and A. Tiomo. 2007. *Risque de Credit: Une Approche Avancée*, Paris: Economica.

[34] Gupton, G., C. Finger, and M. Bhatia. 1997. "Creditmetrics", Technical Report, The Risk Metrics Group.

[35] Hansen, L., and S. Richard. 1987. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." Econometrica 55: 587-613.

[36] Harrison, J., and D. Kreps. 1979. "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory* 20: 381-408.

[37] Hewitt, E., and L. Savage. 1955. "Symmetric Measures on Cartesian Products." *Transaction of the American Mathematical Society* 80: 470-501.

[38] Hull, J., and A. White. 2004. "Valuation of a CDO and an n-to-default CDS Without Monte-Carlo Simulation.“ *Journal of Derivatives* 12: 2, Winter.

[39] Ibragimov, I., and R. Has'minskii. 1981. *Statistical Estimation: Asymptotic Theory.* Berlin: Springer.

[40] Jensen, J. 1995. *Saddlepoint Approximations.* Oxford: Clarendon Press.

[41] Lamb, R., W. Perraudin, and A. Van Landschoot. 2008. "Dynamic Pricing of Synthetic Collateralized Debt Obligations." Working Paper, Imperial College.

[42] Laurent, J.-P., and J. Gregory. 2005. "Basket Default Swaps, CDOs and Factor Copulas." *Journal of Risk* 7: 103-122.

[43] Lee, R., and L. Carter. 1992. "Modeling and Forecasting the Time Series of US Mortality." *Journal of the American Statistical Association* 87: 659-671.

[44] Lindley, D. 1980. "Approximate Bayesian Methods." In J. Bernardo, M. Degroot, D. Lindley, and A. Smith (eds.), *Bayesian Statistics*, Valencia: University Press.

[45] Lucas, A., P. Klaassen, P. Spreij, and S. Straetmans. 2001. "An Analytical Approach to Credit Risk of Large Corporate Bond and Loan Portfolios." *Journal of Banking and Finance* 25: 1635-1644.

[46] Masreliez, C. 1975. "Approximate Non-Gaussian Filtering with Linear State and Observation Relations." *IEEE Transactions on Automatic Control* 20: 107-110.

[47] Robert, C., and G. Casella. 2004. *Monte-Carlo Statistical Methods.* New York, NY: Springer.

[48] Ross, S. 1976. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13: 341-360.

[49] Schloegl, L., and D. O'Kane. 2005. "A Note on the Large Homogeneous Portfolio Approximation with Student-t Copula." *Finance and Stochastics* 9: 577-584.

[50] Schick, I., and S. Mitter. 1994. “Robust Recursive Estimation in the Presence of Heavy-tailed Observation Noise.” *Annals of Statistics* 22: 1045-1080.

[51] Schoenbucher, P. 2001. ”Factor Models for Portfolio Credit Risk When Defaults are Correlated.” *The Journal of Risk Finance* 3: 45-56.

[52] Schoenbucher, P. 2002. ”Taken to the Limit: Simple and Not-so-simple Loan Loss Distributions.” Working Paper, University of Bonn.

[53] Schrager, D. 2006. ”Affine Stochastic Mortality.” *Insurance: Mathematics and Economics* 38: 87-97.

[54] Tierney, L., and J. Kadane. 1986. ”Accurate Approximations for Posterior Moments and Marginal Densities.” *Journal of the American Statistical Association* 81: 82-86.

[55] Vasicek, O. 1987. ”Probability of Loss on Loan Portfolio.” KMV Technical Report.

[56] Vasicek, O. 1991. ”Limiting Loan Loss Probability Distribution.” KMV Technical Report.

Notes

[1]Such derivatives can be traded in organized markets, as for instance some MBS issued by US federal agencies, or Over-the-Counter (OTC). The market sizes at the end of year 2009 were about 30 trillions USD for the Credit Default Swaps (CDS) [Bank for International Settlements (2010)], 9 trillions USD for the MBS, 6.8 trillions USD for the corporate debt and associated credit derivatives, 2.4 trillions USD for the Asset Backed Securities (ABS) backed on credit cards, student loans, etc. (data from the Securities Industry and Financial Market Authority on the website www.sifma.org).

[2]$(y_{i,t})$ denotes the process $y_{i,t}$, $t = 1, 2, \cdots$.

[3]We use the same notation for the dominating measure regardless if $y_{i,t}$ is a continuous variable, and the dominating measure is the Lebesgue measure, or if $y_{i,t}$ is discrete, and $dy_{i,t}$ is the counting measure. This latter situation arises when $y_{i,t}$ is a default indicator with values 0, 1.

[4]But not necessarily systemic risk, that is a risk that may destroy the financial system.

[5]The usual notation in the documents of the Basle Commitee is ρ^2 instead of ρ; we prefer the second notation due to the interpretation of the parameter as a correlation.

[6]Another linear factor specification compatible with affine term structure is $y_{it} = f_t + u_{it}$, $i = 1, ..., n$, where f_t and $u_{i,t}$ are positive variables. However, the conditional distribution of y_{it} given f_t will admit a support (f_t, ∞) depending on the factor value. As seen in Section 2, the large portfolio approximation theorem is valid for a micro-density, which is third-order differentiable with respect to f_t. Thus, it cannot be applied to this type of linear factor model.

[7]A digital CDS is a CDS without legs (and a unitary payment for the contractual recovery rate). Such digital CDS are the standard for short-term CDS. Otherwise, leg and contractual recovery rate adjustments have to be introduced.

[8]In the multiple factor case, a similar approximation as in Proposition 2 can be derived by

replacing $I_{n,t}$ with the matrix $I_{n,t} = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \log h}{\partial f_t \partial f_t'}(y_{i,t}|\hat{f}_{n,t})$ and $I_{n,t}^{-2}K_{n,t}$ with a vector involving among other statistics the third-order derivatives of the cross-sectional log-likelihood w.r.t. the components of f_t.

[9]See Bates (2009), p. 25, for approximations written on the same conditional distribution as our. These approximations are used in the numerical implementation of an algorithm that updates the Laplace transform of the filtering distribution when the joint dynamics of observations and latent states is affine.

[10]Except in the special model of contamination considered in Schick, Mitter (1994).

[11]Except Cipra and Rubio (1991), who take into account a nonlinear measurement equation with additive non-Gaussian noise.

[12]In Lindley (1980), the Laplace approximation for the posterior moments is performed by expanding around the posterior mode, and not around the ML estimate as in Proposition 2. In our setting, this would correspond to perform the expansion around the posterior mode $\tilde{f}_{n,t} = \arg\max_{f_t} \left\{ \sum_{i=1}^{n} \log h(y_{i,t}|f_t) + \log g(f_t|f_{t-1}) \right\}$. It is possible to show that this approach yields the same approximation as Proposition 2 at order $o(1/n)$.

[13]Tierney, Kadane (1986) derive more accurate approximations at order $1/n^2$ by expanding around a modified posterior mode. However, when these more accurate methods are used to approximate a moment $E[\varphi(f_t)|\Omega_t]$ by $E_n[\varphi(f_t)|\Omega_t]$, say, the mapping $\varphi \to E_n[\varphi(f_t)|\Omega_t]$ is not a linear operator [see e.g. formula (A.2) in Tierney, Kadane (1986)]. Thus, these approximations are not appropriate for pricing purposes.

[14]These parameters are such that the variances and correlations between variables $y_{i,t}$ in Example (i) match the variances and correlations between the latent continuous variables $\log(A_{i,t}/L_{i,t})$ in Example (iii).

[15]This model can be extended to a factor with autoregressive dynamics and gamma-type stationary distribution, which corresponds to a discrete-time Cox-Ingersoll-Ross process [see Gouriéroux and Jasiak (2006)].

[16]This payoff is observed at the end of period $(t, t+1)$, not at the beginning of this period.

[17]Similar large portfolio approximations have been proposed e.g. in Lucas et al. (2001), Schonbucher (2002), Frey, McNeil (2003), Schloegl, O'Kane (2005) to approximate the distribution of the percentage portfolio loss in more complex static factor models.

[18]Related large n approximations of the portfolio loss distribution are developed in Gordy (2002) and Dembo, Deuschel, Duffie (2004) based on saddle-point and large deviation techniques, respectively.

[19]The Fourier Transform Inversion formula is presented in Duffie, Pan, Singleton (2000) to compute expectations of indicator and truncated functions of continuous-time affine processes, but is valid for the indicator function of any random variable. It differs from the Fourier Inversion approach used by Laurent, Gregory (2005) since it directly provides the conditional cdf of $\hat{f}_{n,t+1}$ and does not involve the computation of the density of $\hat{f}_{n,t+1}$.

[20]Another one is n_t, $\hat{f}_{nt}$, $\hat{f}_{n,t-1}$.

[21]This formula can also be deduced by expanding the RHS of (3.2) at first-order in $1/n$.

[22]The unconditional default probability is equal to $\Phi\left(-\mu/\sigma\right)$.

[23]The unconditional default correlation between any two firms i and j is given by [see e.g. Gouriéroux, Tiomo (2007), Chapter 7]:

$$corr\left(y_{i,t}, y_{j,t}\right) = \frac{\Phi_2\left(-\mu/\sigma, -\mu/\sigma; \rho\right) - \Phi\left(-\mu/\sigma\right)^2}{\Phi\left(-\mu/\sigma\right)\left[1 - \Phi\left(-\mu/\sigma\right)\right]},$$

where $\Phi_2\left(., .; \rho\right)$ denotes the joint cdf of the bivariate standard Gaussian distribution with correlation coefficient ρ.

[24]Strictly speaking, the CLT cannot be applied to $\hat{f}_{n,t+1} = \frac{1}{n_t}\sum_{i=1}^{n_t} y_{i,t+1}$, since the common factor introduces an equicorrelation structure across the individuals. However, for $\rho = 0.01$ the equicorrelation is weak and the Gaussian approximation implied by the CLT is rather accurate.

[25]The bias reflects the strong negative skewness of the distribution of the relative pricing errors.

Table 1: Statistics for large portfolio approximation (Proposition 2) and granularity adjustment (Proposition 5)

Linear factor model
$\hat{f}_{n,t} = \bar{y}_{n,t},\ I_{n,t} = \frac{1}{\sigma^2},\ \ K_{n,t} = 0,\ I_t = \frac{1}{\sigma^2},\ \beta_t = 0.$

Single-factor stochastic volatility model
$\hat{f}_{n,t} = \frac{1}{n}\sum_{i=1}^{n}(y_{i,t}-\mu)^2,\ I_{n,t} = \frac{1}{2\hat{f}_{n,t}^2},\ \ K_{n,t} = \frac{2}{\hat{f}_{n,t}^3},$ $I_t = \frac{1}{2f_t^2},\ \beta_t = 0.$

SRF model for default correlation
$\hat{f}_{n,t} = \bar{y}_{n,t},\ I_{n,t} = \frac{1}{\bar{y}_{n,t}(1-\bar{y}_{n,t})},\ \ K_{n,t} = 4\frac{1/2-\bar{y}_{n,t}}{\bar{y}_{n,t}^2(1-\bar{y}_{n,t})^2},$ $I_t = \frac{1}{f_t(1-f_t)},\ \ \beta_t = 0.$

Single-factor corporate spread model
$\hat{f}_{n,t} = \Psi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\log y_{i,t} + \log\lambda\right),$ where $\Psi(s) := \frac{d\log\Gamma(s)}{ds},$ $I_{n,t} = \Psi'\left(\hat{f}_{n,t}\right),\ \ K_{n,t} = -\Psi''\left(\hat{f}_{n,t}\right),\ \ I_t = \Psi'(f_t),\ \ \beta_t = \Psi''(f_t).$

Homogenous class of CDS
$\hat{f}_{n,t} = \Psi_\delta^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\left[\delta\log y_{i,t} + (1-\delta)\log(1-y_{i,t})\right]\right)$ where $\Psi_\delta(s) := \delta\Psi(\delta s) + (1-\delta)\Psi((1-\delta)s) - \Psi(s),$ $I_{n,t} = \Psi_\delta'\left(\hat{f}_{n,t}\right),\ \ \ K_{n,t} = -\Psi_\delta''\left(\hat{f}_{n,t}\right),$ $I_t = \Psi_\delta'(f_t),\ \ \ \beta_t = \Psi_\delta''(f_t).$

Table 2: Relative pricing errors

$\rho = 0.01$	$n = 100$				$n = 1000$			
α	.025	.05	.10	.125	.025	.05	.10	.125
Mean					.0254	.0068	price below tick	
5% Qu					-.0497	-.5362	price below tick	
25% Qu					-.0141	-.2179	price below tick	
Median	approximation breaks down				.0029	.0009	price below tick	
75 % Qu					.0320	.2183	price below tick	
95% Qu					.1808	.5757	price below tick	
Median abs					.0201	.2179	price below tick	

$\rho = 0.10$	$n = 100$				$n = 1000$			
α	.025	.05	.10	.125	.025	.05	.10	.125
Mean	.0260	.0287	-.0157	-.0190	.0004	-.0023	-.0004	-.0109
5% Qu	-.4735	-.8361	-1.1743	-1.3239	-.0462	-.0983	-.2215	-.2551
25% Qu	-.0728	-.1484	-.2799	-.3652	-.0125	-.0294	-.0770	-.0908
Median	.0033	.0057	-.0055	-.0206	.0009	.0013	-.0034	-.0051
75% Qu	.0887	.1942	.2682	.3064	.0123	.0293	.0739	.1018
95% Qu	.5385	.7743	.8887	1.1067	.0426	.0831	.2284	.3171
Median abs	.0792	.1662	.2718	.3363	.0124	.0294	.0754	.0960

$\rho = 0.30$	$n = 100$				$n = 1000$			
α	.025	.05	.10	.125	.025	.05	.10	.125
Mean	-.0573	-.0750	-.0931	-.0975	-.0065	-.0076	-.0084	-.0086
5% Qu	-.4988	-.6628	-.8342	-.8837	-.0722	-.1000	-.1582	-.2083
25% Qu	-.0611	-.0800	-.1120	.1262	-.0180	-.0280	-.0518	-.0631
Median	-.0027	-.0029	-.0006	.0007	-.0000	-.0016	-.0062	-.0085
75% Qu	.0240	.0381	.0626	.0758	.0143	.0235	.0440	.0504
95% Qu	.0925	.1408	.2103	.2463	.0486	.0749	.1300	.1723
Median abs	.0358	.0505	.0761	.0888	.0159	.0257	.0484	.0564

Figure 1: Prediction interval as a function of the available information

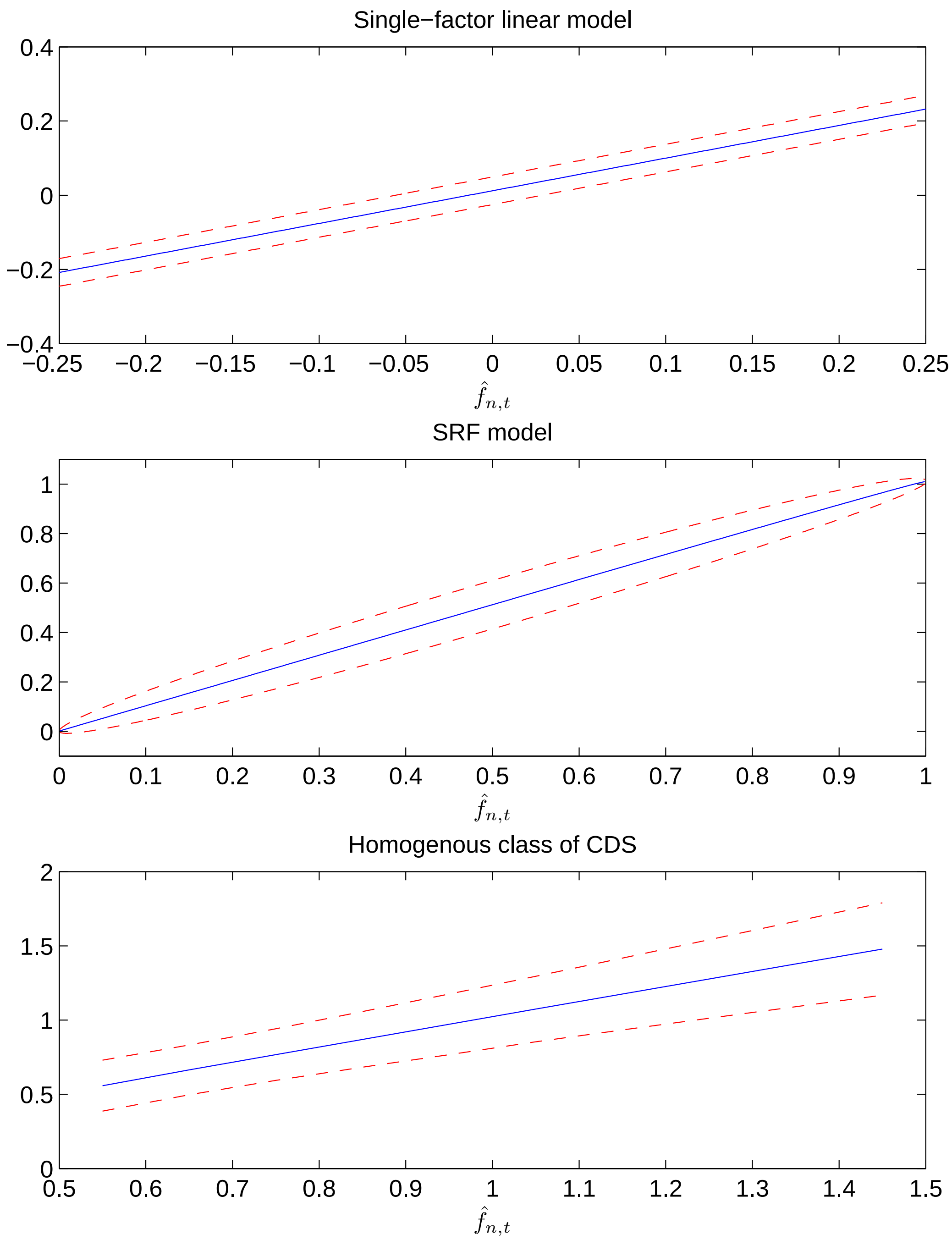


Figure 2: Transition density of the factor

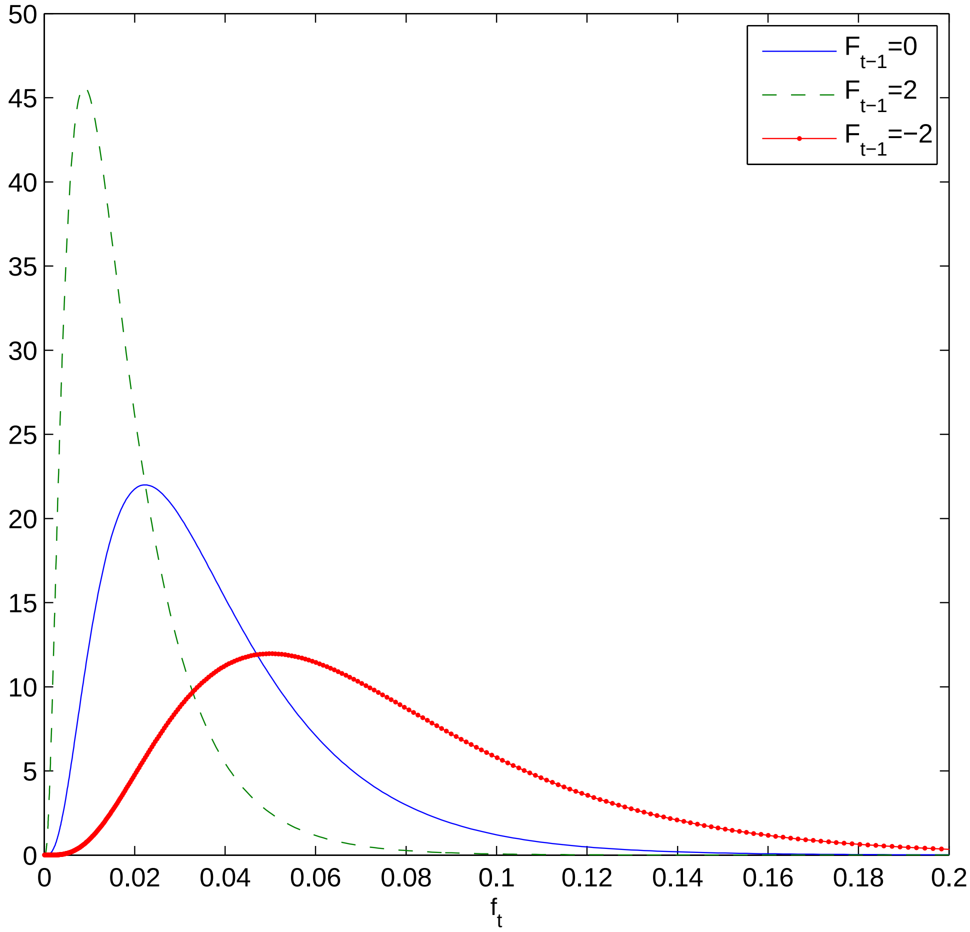


Figure 3: Predictive distribution of the factor given investors' information

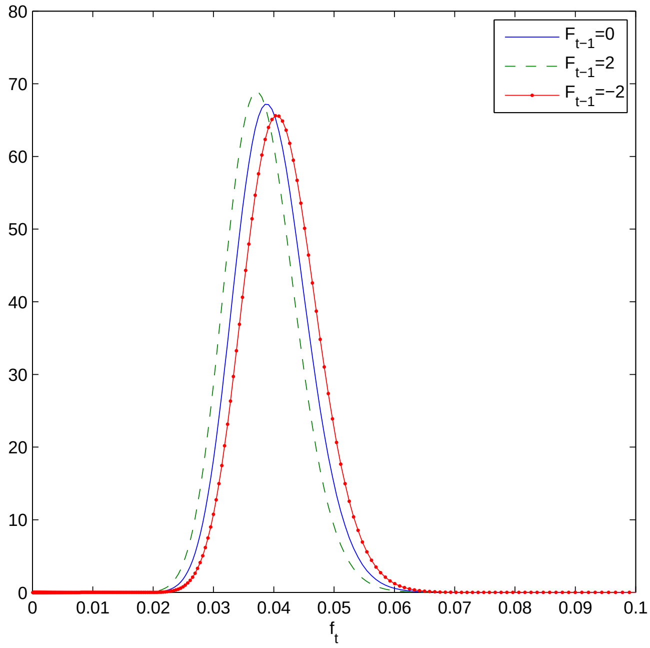


Figure 4: Gaussian approximation of the factor predictive density

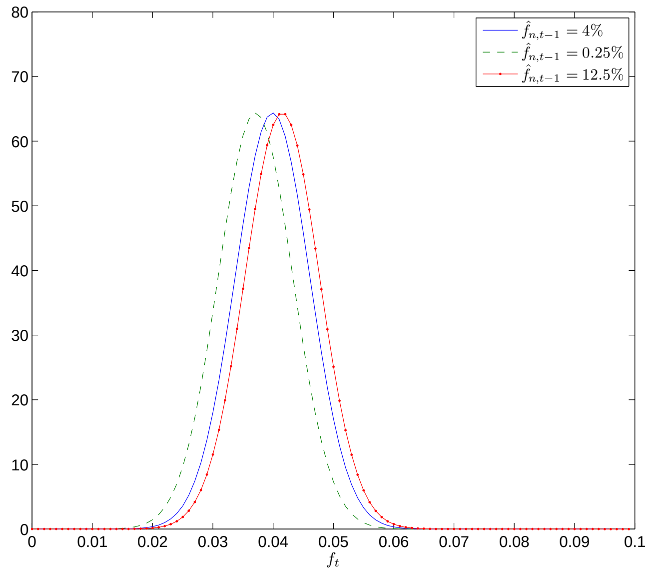


Figure 5: True price of the α-to-default swap for different values of the asset correlation

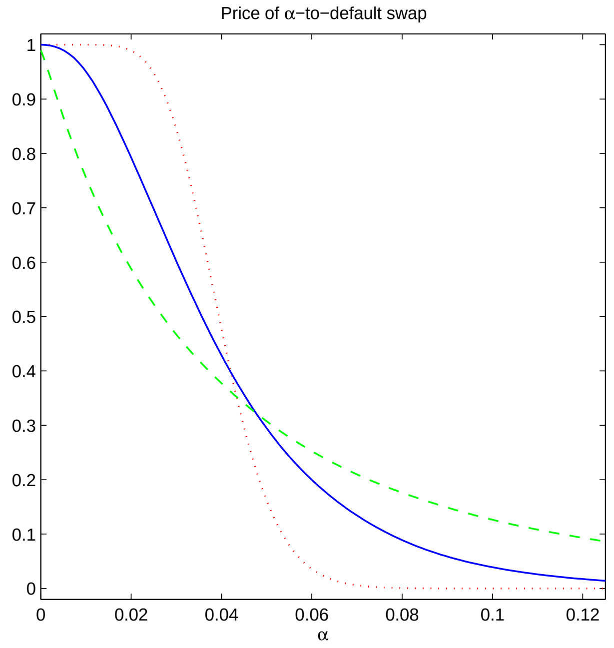


Figure 6: True price of the α-to-default swap for different values of the lagged factor

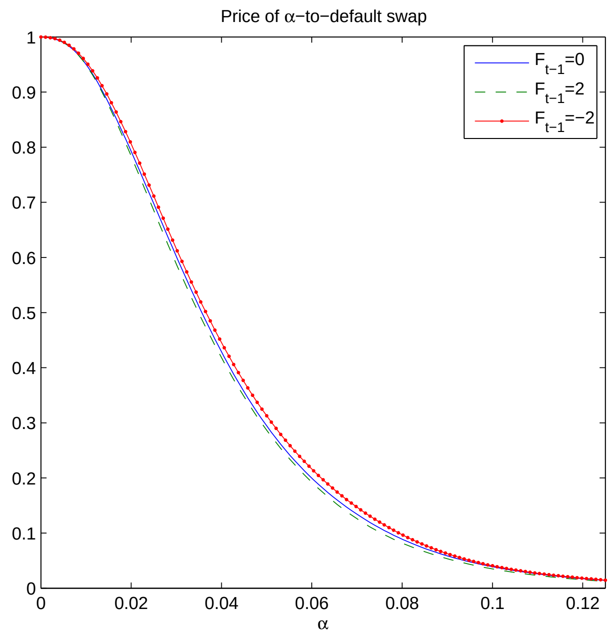


Figure 7: Approximate price of the α-to-default swap for different values of asset correlation

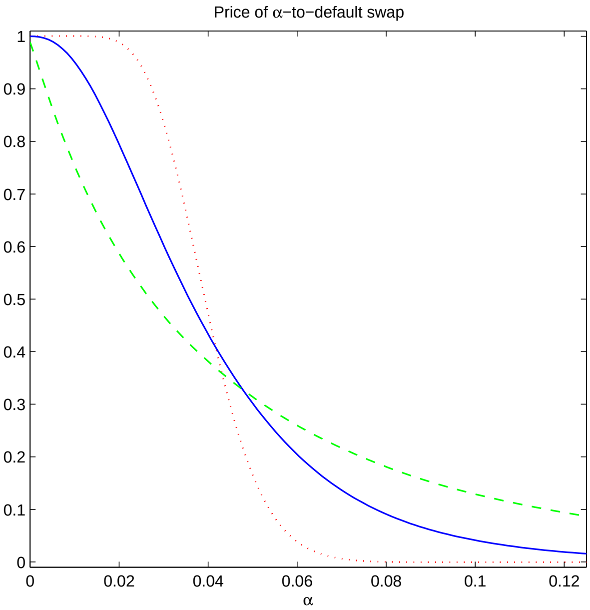
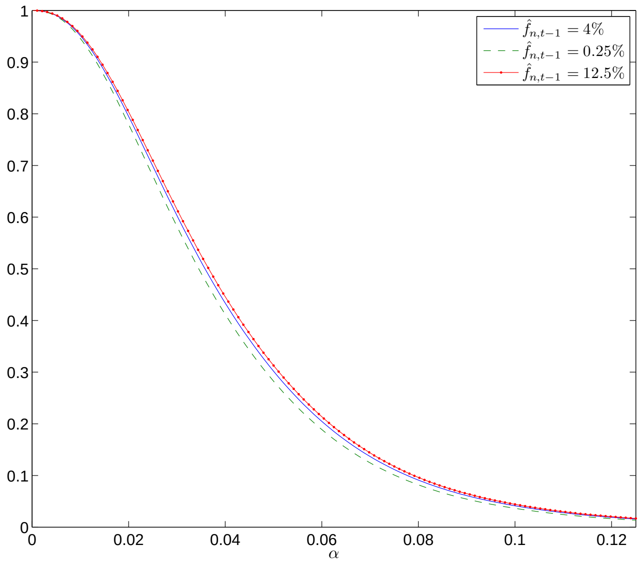


Figure 8: Approximate price of the α-to-default swap for different values of the lagged factor



APPENDIX 1: Proof of Proposition 2

i) Let us first derive an approximation for the conditional Laplace transform of f_t given $\underline{y_{1,t}}, ..., \underline{y_{n,t}}$ and $\underline{f_{t-1}}$:

$$\mathcal{L}_{nt}(u) = E\left[\exp(uf_t)\,|\underline{y_{1,t}}, ..., \underline{y_{n,t}}, \underline{f_{t-1}}\right] = \frac{\displaystyle\int e^{uf_t} g(f_t|f_{t-1}) \prod_{i=1}^{n} h(y_{i,t}|f_t)\, df_t}{\displaystyle\int g(f_t|f_{t-1}) \prod_{i=1}^{n} h(y_{i,t}|f_t)\, df_t}, \tag{A.1}$$

which depends only on $y_{1,t}, ..., y_{nt}$ and f_{t-1}.

Let us expand the micro-density around $\hat{f}_{nt}$:

$$\begin{aligned}
\sum_{i=1}^{n} \log h(y_{i,t}|f_t) &= \sum_{i=1}^{n} \log h\left(y_{i,t}|\hat{f}_{nt}\right) \\
&+ \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \log h}{\partial f_t^2}\left(y_{i,t}|\hat{f}_{nt}\right)\left[\sqrt{n}\left(f_t - \hat{f}_{nt}\right)\right]^2 \\
&+ \frac{1}{6\sqrt{n}}\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^3 \log h}{\partial f_t^3}\left(y_{i,t}|\hat{f}_{nt}\right)\left[\sqrt{n}\left(f_t - \hat{f}_{nt}\right)\right]^3 \\
&+ \frac{1}{24n}\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^4 \log h}{\partial f_t^4}\left(y_{i,t}|\hat{f}_{nt}\right)\left[\sqrt{n}\left(f_t - \hat{f}_{nt}\right)\right]^4 + o(1/n).
\end{aligned}$$

Let us introduce the change of variable:

$$X = I_{nt}^{1/2}\sqrt{n}\left(f_t - \hat{f}_{nt}\right) \iff f_t = \hat{f}_{nt} + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} X.$$

Then, we have:

$$\sum_{i=1}^{n} \log h(y_{i,t}|f_t) = \sum_{i=1}^{n} \log h\left(y_{i,t}|\hat{f}_{nt}\right) - \frac{1}{2}X^2 + \frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} Q_{nt} X^4 + o(1/n),$$

where:

$$J_{nt} = I_{nt}^{-3/2} K_{nt} \text{ and } Q_{nt} = I_{nt}^{-2}\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^4 \log h}{\partial f_t^4}\left(y_{i,t}|\hat{f}_{nt}\right).$$

Thus:

$$\begin{aligned}\prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) &= \prod_{i=1}^{n} h\left(y_{i,t}|\hat{f}_{nt}\right)\exp\left(-\frac{1}{2}X^2\right)\exp\left(\frac{1}{6\sqrt{n}}J_{nt}X^3+\frac{1}{24n}Q_{nt}X^4+o(1/n)\right)\\ &= \prod_{i=1}^{n} h\left(y_{i,t}|\hat{f}_{nt}\right)\exp\left(-\frac{1}{2}X^2\right)\\ &\quad \left[1+\frac{1}{6\sqrt{n}}J_{nt}X^3+\frac{1}{24n}Q_{nt}X^4+\frac{1}{72n}J_{nt}^2X^6+o(1/n)\right]. \end{aligned} \tag{A.2}$$

Similarly, we have an expansion for $\log g(f_t|f_{t-1})$ as:

$$\begin{aligned}\log g(f_t|f_{t-1}) &= \log g\left(\hat{f}_{nt}+\frac{1}{\sqrt{n}}I_{nt}^{-1/2}X|f_{t-1}\right)\\ &= \log g\left(\hat{f}_{nt}|f_{t-1}\right)+\frac{1}{\sqrt{n}}I_{nt}^{-1/2}A_{nt}X+\frac{1}{2n}I_{nt}^{-1}B_{nt}X^2+o(1/n),\end{aligned}$$

where:

$$A_{nt}=\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|f_{t-1}\right) \quad \text{and} \quad B_{nt}=\frac{\partial^2 \log g}{\partial f_t^2}\left(\hat{f}_{nt}|f_{t-1}\right).$$

Thus:

$$\begin{aligned}g(f_t|f_{t-1}) &= g\left(\hat{f}_{nt}|f_{t-1}\right)\exp\left(\frac{1}{\sqrt{n}}I_{nt}^{-1/2}A_{nt}X+\frac{1}{2n}I_{nt}^{-1}B_{nt}X^2+o(1/n)\right)\\ &= g\left(\hat{f}_{nt}|f_{t-1}\right)\left[1+\frac{1}{\sqrt{n}}I_{nt}^{-1/2}A_{nt}X+\frac{1}{2n}I_{nt}^{-1}B_{nt}X^2+\frac{1}{2n}I_{nt}^{-1}A_{nt}^2X^2+o(1/n)\right].\end{aligned} \tag{A.3}$$

Finally, we have an expansion for $\exp\left(uf_t\right)$:

$$\begin{aligned}\exp\left(uf_t\right) &= \exp\left(u\hat{f}_{nt}\right)\exp\left(\frac{u}{\sqrt{n}}I_{nt}^{-1/2}X\right)\\ &= \exp\left(u\hat{f}_{nt}\right)\left[1+\frac{u}{\sqrt{n}}I_{nt}^{-1/2}X+\frac{u^2}{2n}I_{nt}^{-1}X^2+o(1/n)\right].\end{aligned} \tag{A.4}$$

Let us now substitute expansions (A.2)-(A.4) into the numerator in equation (A.1) (the denominator is obtained by setting $u=0$). We have:

$$\begin{aligned}
&\int e^{uf_t} g(f_t|f_{t-1}) \prod_{i=1}^{n} h\left(y_{i,t}|f_t\right) df_t = e^{u\hat{f}_{nt}} \prod_{i=1}^{n} h\left(y_{i,t}|\hat{f}_{nt}\right) g\left(\hat{f}_{nt}|f_{t-1}\right) \\
&\quad E_X\left[\left(1+\frac{u}{\sqrt{n}} I_{nt}^{-1/2} X + \frac{u^2}{2n} I_{nt}^{-1} X^2 + o(1/n)\right)\right. \\
&\qquad \left(1+\frac{1}{\sqrt{n}} I_{nt}^{-1/2} A_{nt} X + \frac{1}{2n} I_{nt}^{-1}\left(B_{nt}+A_{nt}^2\right) X^2 + o(1/n)\right) \\
&\qquad\quad \left.\left(1+\frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} Q_{nt} X^4 + \frac{1}{72n} J_{nt}^2 X^6 + o(1/n)\right)\right],
\end{aligned}$$

where the expectation E_X is w.r.t. the standard normal variable X. Since odd power moments of X are equal to zero, the terms of order $1/\sqrt{n}$ [and similarly the terms of order $1/\left(n\sqrt{n}\right)$, if the expansion is considered up to order $1/n^2$] cancel and the expectation is equal to:

$$1+\frac{u}{n}\left[I_{nt}^{-1} A_{nt} + \frac{1}{2} I_{nt}^{-1/2} J_{nt}\right] + \frac{1}{2n} u^2 I_{nt}^{-1} + \Lambda_{nt} + O(1/n^2),$$

where:

$$\Lambda_{nt} = \frac{1}{2n} I_{nt}^{-1}\left(B_{nt}+A_{nt}^2\right) + \frac{1}{2n} I_{nt}^{-1/2} J_{nt} A_{nt} + \frac{1}{8n} Q_{nt} + \frac{1}{72n} J_{nt}^2 E\left[X^6\right],$$

is independent of u. Thus, we deduce:

$$\begin{aligned}
\mathcal{L}_{nt}(u) &= e^{u\hat{f}_{nt}} \frac{1+\frac{u}{n}\left(I_{nt}^{-1} A_{nt} + \frac{1}{2} I_{nt}^{-1/2} J_{nt}\right) + \frac{1}{2n} u^2 I_{nt}^{-1} + \Lambda_{nt} + O(1/n^2)}{1+\Lambda_{nt}+O(1/n^2)} \\
&= e^{u\hat{f}_{nt}} \left(1+\frac{u}{n}\left(I_{nt}^{-1} A_{nt} + \frac{1}{2} I_{nt}^{-1/2} J_{nt}\right) + \frac{1}{2n} u^2 I_{nt}^{-1} + \Lambda_{nt} + O(1/n^2)\right)\left(1-\Lambda_{nt}+O(1/n^2)\right) \\
&= e^{u\hat{f}_{nt}} \left[1+\frac{u}{n}\left(I_{nt}^{-1} A_{nt} + \frac{1}{2} I_{nt}^{-1/2} J_{nt}\right) + \frac{1}{2n} u^2 I_{nt}^{-1} + O(1/n^2)\right].
\end{aligned}$$

By definition of J_{nt} and A_{nt}, we conclude:

$$\mathcal{L}_{nt}(u) = e^{u\hat{f}_{nt}}\left\{1+\frac{u}{n}\left(I_{nt}^{-1}\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|f_{t-1}\right) + \frac{1}{2} I_{nt}^{-2} K_{nt}\right) + \frac{1}{2n} u^2 I_{nt}^{-1}\right\} + O(1/n^2),$$

and

$$\mathcal{L}_{nt}(u) = \exp\left\{u\hat{f}_{nt} + \frac{u}{n}\left(I_{nt}^{-1}\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|f_{t-1}\right) + \frac{1}{2}I_{nt}^{-2}K_{nt}\right) + \frac{1}{2n}u^2 I_{nt}^{-1} + O(1/n^2)\right\}.$$

ii) Another approximation valid at order $1/n$ can be obtained by replacing f_{t-1} by $\hat{f}_{n,t-1}$. We have:

$$\mathcal{L}_{nt}(u) = \exp\left\{u\hat{f}_{nt} + \frac{u}{n}\left(I_{nt}^{-1}\frac{\partial \log g}{\partial f_t}\left(\hat{f}_{nt}|\hat{f}_{n,t-1}\right) + \frac{1}{2}I_{nt}^{-2}K_{nt}\right) + \frac{1}{2n}u^2 I_{nt}^{-1} + o(1/n)\right\}.$$

Then, Proposition 2 follows.

APPENDIX 2: Examples

In this Appendix we derive the expressions of the summary statistics $\hat{f}_{nt}$, I_{nt}, K_{nt}, I_t and β_t for Example (v) (the derivation of the statistics for the other examples given in Table 1 is similar and available from the authors on request).

We have:

$$\begin{aligned}
\frac{\partial \log h(y_{i,t}|f_t)}{\partial f_t} &= -\Psi_\delta\left(f_t\right) + \delta \log y_{i,t} + (1-\delta)\log\left(1-y_{i,t}\right), \\
\frac{\partial^2 \log h(y_{i,t}|f_t)}{\partial f_t^2} &= -\Psi_\delta^{'}\left(f_t\right), \qquad \frac{\partial^3 \log h(y_{i,t}|f_t)}{\partial f_t^3} = -\Psi_\delta^{''}\left(f_t\right),
\end{aligned}$$

where $\Psi_\delta\left(f_t\right) := \delta\Psi\left(\delta f_t\right) + (1-\delta)\Psi\left((1-\delta)f_t\right) - \Psi\left(f_t\right)$ and $\Psi(s) := \dfrac{d\log\Gamma(s)}{ds}$. We deduce that $\hat{f}_{n,t}$ solves the equation $\Psi_\delta(\hat{f}_{n,t}) = \dfrac{1}{n}\sum_{i=1}^{n}[\delta\log y_{i,t} + (1-\delta)\log(1-y_{i,t})]$, $I_{n,t} = \Psi_\delta^{'}\left(\hat{f}_{n,t}\right)$, $K_{n,t} = -\Psi_\delta^{''}\left(\hat{f}_{n,t}\right)$, and $I_t = \Psi_\delta^{'}\left(f_t\right)$. Since I_t is a conditional variance, we have $I_t > 0$, which implies that $\Psi_\delta^{'}(.)$ is monotone increasing. Thus, $\hat{f}_{n,t} = \Psi_\delta^{-1}\left(\dfrac{1}{n}\sum_{i=1}^{n}[\delta\log y_{i,t} + (1-\delta)\log(1-y_{i,t})]\right)$.

Finally, let us derive β_t. We have:

$$\begin{aligned}\beta_t &= Cov\left(\frac{\partial \log h(y_{i,t}|f_t)}{\partial f_t}, \frac{\partial^2 \log h(y_{i,t}|f_t)}{\partial f_t^2} + \left(\frac{\partial \log h(y_{i,t}|f_t)}{\partial f_t}\right)^2 |f_t\right) = E\left[\left(\frac{\partial \log h(y_{i,t}|f_t)}{\partial f_t}\right)^3 |f_t\right] \\ &= E\left[\left(Z_{i,t} - E\left[Z_{i,t}|f_t\right]\right)^3 |f_t\right] = M_t^{'''}(0),\end{aligned}$$

where $Z_{i,t} := \delta \log y_{i,t} + (1-\delta)\log(1-y_{i,t})$ and $M_t(s) := \log E\left[\exp(sZ_{i,t})|f_t\right]$. We have:

$$\begin{aligned}E\left[\exp(sZ_{i,t})|f_t\right] &= E\left[y_{i,t}^{\delta s}(1-y_{i,t})^{(1-\delta)s}|f_t\right] \\ &= \frac{\Gamma(f_t)}{\Gamma(\delta f_t)\Gamma((1-\delta)f_t)}\int_0^1 y_{i,t}^{\delta(f_t+s)-1}(1-y_{i,t})^{(1-\delta)(f_t+s)-1}dy_{i,t} \\ &= \frac{\Gamma(f_t)}{\Gamma(\delta f_t)\Gamma((1-\delta)f_t)}\frac{\Gamma(\delta(f_t+s))\Gamma((1-\delta)(f_t+s))}{\Gamma(f_t+s)}.\end{aligned}$$

Thus:

$$\beta_t = M_t^{'''}(0) = \delta^3\Psi''(\delta f_t) + (1-\delta)^3\Psi''((1-\delta)f_t) - \Psi''(f_t) = \Psi_\delta''(f_t).$$

APPENDIX 3: Proof of Proposition 5

The proof of Proposition 5 relies on a higher-order stochastic expansion for $\hat{f}_{n,t+\tau}$. This stochastic expansion is derived from results in ML theory that are recalled below.

i) A useful result in ML theory

Let $\hat{\theta}_n = \arg\max_\theta \sum_{i=1}^n \log h(y_i;\theta)$ be the ML estimator of parameter θ in the p.d.f. $h(y;\theta)$. The asymptotic expansion of $\hat{\theta}_n$ at order $1/n$ is given by [see Gouriéroux, Monfort (1995), Chapter 23]:

$$\hat{\theta}_n - \theta = \frac{1}{\sqrt{n}}I^{-1}A_n + \frac{1}{n}\left[I^{-2}A_nB_n + \frac{1}{2}I^{-3}KA_n^2\right] + o_p(1/n), \quad \text{(A.5)}$$

where:

$$I = E\left[-\frac{\partial^2 \log h\,(y_i;\theta)}{\partial\theta^2}\right] \quad , \quad K = E\left[\frac{\partial^3 \log h\,(y_i;\theta)}{\partial\theta^3}\right],$$

and:

$$A_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial \log h}{\partial\theta}\,(y_i;\theta) \ , \quad B_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\partial^2 \log h\,(y_i;\theta)}{\partial\theta^2} + I\right].$$

The quantity K can be rewritten in terms of covariances of the first- and second-order derivatives of the log density. More precisely we have:

$$\frac{\partial \log h}{\partial\theta} = \frac{1}{h}\frac{\partial h}{\partial\theta}, \tag{A.6}$$

$$\begin{aligned}\frac{\partial^2 \log h}{\partial\theta^2} &= \frac{1}{h}\frac{\partial^2 h}{\partial\theta^2} - \frac{1}{h^2}\left(\frac{\partial h}{\partial\theta}\right)^2 \\ &= \frac{1}{h}\frac{\partial^2 h}{\partial\theta^2} - \left(\frac{\partial \log h}{\partial\theta}\right)^2,\end{aligned} \tag{A.7}$$

and:

$$\begin{aligned}\frac{\partial^3 \log h}{\partial\theta^3} &= \frac{1}{h}\frac{\partial^3 h}{\partial\theta^3} - \frac{3}{h^2}\frac{\partial h}{\partial\theta}\frac{\partial^2 h}{\partial\theta^2} + \frac{2}{h^3}\left(\frac{\partial h}{\partial\theta}\right)^3 \\ &= \frac{1}{h}\frac{\partial^3 h}{\partial\theta^3} - 3\frac{\partial \log h}{\partial\theta}\frac{\partial^2 \log h}{\partial\theta^2} - \left(\frac{\partial \log h}{\partial\theta}\right)^3.\end{aligned} \tag{A.8}$$

By taking the expectation on both sides of equations (A.6) and (A.7), we get $E\left[\frac{\partial \log h(y_i;\theta)}{\partial\theta}\right] = 0$ and $I = E\left[\left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}\right)^2\right]$, respectively. By taking the expectation on both sides of equation (A.8), we get:

$$\begin{aligned}K &= -3Cov\left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}, \frac{\partial^2 \log h(y_i;\theta)}{\partial\theta^2}\right) - E\left[\left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}\right)^3\right] \\ &= -2Cov\left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}, \frac{\partial^2 \log h(y_i;\theta)}{\partial\theta^2}\right) \\ &\quad -Cov\left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}, \frac{\partial^2 \log h(y_i;\theta)}{\partial\theta^2} + \left(\frac{\partial \log h(y_i;\theta)}{\partial\theta}\right)^2\right).\end{aligned} \tag{A.9}$$

ii) Stochastic expansion of $\hat{f}_{n,t+\tau}$

The asymptotic expansion of the ML estimator $\hat{f}_{n,t+\tau} = \arg\max_{f_{t+\tau}} \sum_{i=1}^{n} \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)$ at order $1/n$ is derived from equation (A.5), by replacing the parameter θ with the factor value $f_{t+\tau}$ and by computing the expectations conditional on $f_{t+\tau}$. The stochastic expansion is given by:

$$\hat{f}_{n,t+\tau} - f_{t+\tau} = \frac{1}{\sqrt{n}} I_{t+\tau}^{-1} A_{n,t+\tau} + \frac{1}{n}\left[I_{t+\tau}^{-2} A_{n,t+\tau} B_{n,t+\tau} + \frac{1}{2} I_{t+\tau}^{-3} K_{t+\tau} A_{n,t+\tau}^{2}\right] + o_p(1/n), \quad \text{(A.10)}$$

where:

$$I_{t+\tau} = E\left[-\frac{\partial^2 \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f^2}|f_{t+\tau}\right] \quad , \quad K_{t+\tau} = E\left[\frac{\partial^3 \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f^3}|f_{t+\tau}\right],$$

and:

$$A_{n,t+\tau} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log h}{\partial f}\left(y_{i,t+\tau}|f_{t+\tau}\right) \ , \ \ B_{n,t+\tau} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{\partial^2 \log h\left(y_{i,t+\tau}|f_{t+\tau}\right)}{\partial f^2} + I_{t+\tau}\right].$$

Conditionally on Ω_t and (f_t), the random vector $C_{n,t+\tau} := \left(A_{n,t+\tau}, B_{n,t+\tau}\right)'$ is such that:

$$E\left[C_{n,t+\tau}|\Omega_t, (f_t)\right] = 0, \quad \text{(A.11)}$$

and:

$$V\left[C_{n,t+\tau}|\Omega_t, (f_t)\right] = \begin{bmatrix} I_{t+\tau} & W_{t+\tau} \\ W_{t+\tau} & R_{t+\tau} \end{bmatrix}, \quad \text{(A.12)}$$

where $W_{t+\tau} = Cov\left(\frac{\partial \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f}, \frac{\partial^2 \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f^2}|f_{t+\tau}\right)$ and $R_{t+\tau} = V\left[\frac{\partial^2 \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f^2}|f_{t+\tau}\right]$.

Moreover from (A.9) we have:

$$K_{t+\tau} = -2W_{t+\tau} - \beta_{t+\tau}, \quad \text{(A.13)}$$

where $\beta_{t+\tau} = Cov\left(\frac{\partial \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f}, \frac{\partial^2 \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f^2} + \left(\frac{\partial \log h(y_{i,t+\tau}|f_{t+\tau})}{\partial f}\right)^2 |f_{t+\tau}\right)$.

iii) Asymptotic expansion of the derivative price

The price at t of the derivative with payoff $\exp\left(u\hat{f}_{n,t+\tau}\right)$ is given by:

$$\begin{aligned}\tilde{\pi}_{n,t}(u,\tau) &= E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})e^{u\hat{f}_{n,t+\tau}}|\Omega_t\right] \\ &= E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})E\left[e^{u\hat{f}_{n,t+\tau}}|\Omega_t,(f_t)\right]|\Omega_t\right].\end{aligned}$$

By using stochastic expansion (A.10), we have:

$$\begin{aligned}&E\left[e^{u\hat{f}_{n,t+\tau}}|\Omega_t,(f_t)\right] \\ =\;& e^{uf_{t+\tau}}E\left[\exp\left(\frac{u}{\sqrt{n}}I_{t+\tau}^{-1}A_{n,t+\tau}+\frac{u}{n}\left[I_{t+\tau}^{-2}A_{n,t+\tau}B_{n,t+\tau}+\frac{1}{2}I_{t+\tau}^{-3}K_{t+\tau}A_{n,t+\tau}^2\right]\right)|(f_t)\right]+o(1/n).\end{aligned}$$

By expanding the exponential function, and using (A.11)-(A.13), we get:

$$\begin{aligned}E\left[e^{u\hat{f}_{n,t+\tau}}|\Omega_t,(f_t)\right] &= e^{uf_{t+\tau}}\left(1+\frac{u}{n}I_{t+\tau}^{-2}W_{t+\tau}+\frac{u}{2n}I_{t+\tau}^{-2}K_{t+\tau}+\frac{u^2}{2n}I_{t+\tau}^{-1}\right)+o(1/n) \\ &= e^{uf_{t+\tau}}\left(1+\frac{u^2}{2n}I_{t+\tau}^{-1}-\frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau}\right)+o(1/n) \\ &= \exp\left(uf_{t+\tau}+\frac{u^2}{2n}I_{t+\tau}^{-1}-\frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau}\right)+o(1/n).\end{aligned}$$

We conclude by using the iterated expectation theorem and the Markov property of process (f_t):

$$\begin{aligned}&\tilde{\pi}_{n,t}(u,\tau) \\ =\;& E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})\exp\left(uf_{t+\tau}+\frac{u^2}{2n}I_{t+\tau}^{-1}-\frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau}\right)|\Omega_t\right]+o(1/n) \\ =\;& E\left[E\left[m(f_t)m(f_{t+1})\cdots m(f_{t+\tau-1})\exp\left(uf_{t+\tau}+\frac{u^2}{2n}I_{t+\tau}^{-1}-\frac{u}{2n}I_{t+\tau}^{-2}\beta_{t+\tau}\right)|f_t\right]|\Omega_t\right]+o(1/n).\end{aligned}$$

APPENDIX 4: Numerical illustration to basket default derivatives

i) Computation by simulation of the true price of exponential derivatives

The true price $\tilde{\pi}_{n,t}(u,1)$ in (3.2) can be computed with Monte-Carlo simulation by drawing first a sample $\{f_t^s\}_{s=1,...,S}$ from the density of f_t given Ω_t, and then drawing a value f_{t+1}^s from the density of f_{t+1} given f_t^s for any $s = 1, ..., S$. To simulate from the density of f_t given Ω_t we can use the acceptance-rejection algorithm [e.g., Robert, Casella (2004)]. We have from (3.4):

$$g(f_t|\Omega_t) \propto (f_t)^{n_{t-1}-n_t} (1-f_t)^{n_t} g(f_t|f_{t-1}),$$

where $\propto$ denotes equality of the densities up to a scale factor (depending on the conditioning variables only). By definition of the ML estimator $\hat{f}_{n,t}$, we have:

$$(f_t)^{n_{t-1}-n_t} (1-f_t)^{n_t} \leq \left(\hat{f}_{n,t}\right)^{n_{t-1}-n_t} \left(1-\hat{f}_{n,t}\right)^{n_t}, \text{ for any value of } f_t.$$

Thus, the density of f_t given Ω_t is upper bounded, and the density of f_t given f_{t-1} is a majorizing density. The acceptance-rejection algorithm to draw f_t^s works as follows:

i) Generate a random draw $\tilde{f}_t$ from the density of f_t given f_{t-1}.

ii) Generate a uniform variable $U \sim \mathcal{U}(0,1)$, independent of $\tilde{f}_t$.

iii) If $U \leq \dfrac{\left(\tilde{f}_t\right)^{n_{t-1}-n_t}\left(1-\tilde{f}_t\right)^{n_t}}{\left(\hat{f}_{n,t}\right)^{n_{t-1}-n_t}\left(1-\hat{f}_{n,t}\right)^{n_t}}$, then $f_t^s = \tilde{f}_t$. Otherwise, return to i).

ii) Computation by simulation of the approximate price of exponential derivatives

The integral in (3.6) can be computed by Monte-Carlo simulation. The simulated sample of future factor values is obtained by first drawing f_t^s from the Gaussian distribution $\hat{g}_{nt}$, and then drawing f_{t+1}^s from the distribution of f_{t+1} given f_t^s, for $s = 1, \cdots, S$, where the number of replications S is large. The integral in the RHS of equation (3.6) is computed by averaging the values $\left\{1 + [\exp(u/n_t) - 1] f_{t+1}^s\right\}^{n_t}$ for $s = 1, \cdots, S$.